SOME PROBLEMS AND RESULTS IN NUMBER THEORY

P. Erdös

Hungarian Academy of Sciences

During my very long life I published many papers of similar title. Here I want to discuss some of my favorite problems many of which go back 50 years and which I hope are still alive and will outlive me.

Recently Graham and I published a book entitled "Old and new problems and results in combinatorial number theory" Monographie N°28 de L'Enseignement Mathématique, Univ. de Genève. This booklet contains many references and I do not repeat them here. I will refer to this book as I.

I gave a talk in January 1984 at a meeting organized by K. Alladi of Matscience, Madras, held at Ooty. The proceedings of that conference will also appear as the lecture note series of Springer Verlag and I will try to minimize the overlap between the two papers. This is the reason why I will not discuss problems on covering congruences in detail.

First I discuss some of my problems which have been solved in the last two years. I hope the reader will forgive a very old man if he inserts a few historical remarks.

Soon two further papers of mine on related topics will appear. The first one is in a volume dedicated to the memory of E. Straus and was a lecture given at the number theory meeting at the University of Alberta, April 28-30, 1983. The second was given at the meeting on number theory at Luminy late in May 1983. Of course I cannot entirely avoid some overlap. Richard Guy recently published a very nice book on problems in number theory.

I

1. In a paper written jointly with L. Mirsky we investigated among others the following question: Let $F(x)$ be the greatest integer k so that there exists a run of k consecutive integers $n+1, n+2, \dots, n+k$, $n+k \leq x$ for which all the values $d(n+i)$, $1 \leq i \leq k$, are distinct. We proved $F(x) > c\,\frac{(\log x)^{1/2}}{\log\log x}$ and with some more work we could have proved $F(x) > c\,(\log x)^{1-\varepsilon}$. Our method certainly could not give $F(x) > c\,\frac{\log x}{\log\log x}$. From above we could not prove anything better than $F(x) < \exp\left(\frac{c(\log x)^{1/2}}{\log\log x}\right)$. We conjectured that the true order of magnitude of $F(x)$ is $(\log x)^c$. We also stated a related problem: What is the longest run of consecutive integers not exceeding x all of which have the same number of divisors? This problem seemed very difficult to us and we even considered it hopeless to prove that $d(n) = d(n+1)$ has infinitely many solutions. This latter problem was generally considered to be very difficult. To my great surprise Claudia Spiro proved about 4 years ago in her thesis that

$$d(n) = d(n+5040)$$

has infinitely many solutions. A few weeks ago Heath-Brown, using Spiro's method and some further ideas, completely solved our problem. He in fact proved that the number of integers $n < x$ for which

$$d(n) = d(n+1) \tag{1}$$

is at least $cx/(\log x)^4$. He remarks that his method cannot deal with $2d(n) = d(n+1)$ and of course it does not deal with $d(n) = d(n+1) = d(n+2)$.

In a forthcoming paper of Pomerance, Sárközy and myself we prove that

(1) has at most $cx/(\log\log x)^{1/2}$ solutions which is probably the correct order of magnitude.

The same problem can be asked for other number theoretic functions e.g. for $\varphi(n)$. Presumably for every k $\varphi(n) = \varphi(n+1) = \cdots = \varphi(n+k)$ has infinitely many solutions, but this seems unattackable even for $k = 1$.*) I further expect that for every c and $x > x_0(c)$ there are $(\log x)^c$ consecutive integers not exceeding x for which all the values $\varphi(n+i)$, $1 \leq i < (\log x)^c$, are distinct. Perhaps this will not be very difficult but I certainly have not been able to prove it. In a forthcoming paper Pomerance, Sárközy and I prove that if the integers $\varphi(n+i)$, $1 \leq i \leq k_n$ are all distinct then

$$k_n < n \exp(-(\log n)^{1/3}) .$$

The true order of magnitude of k_n is probably $O(n^\varepsilon)$.

2. Denote by $1 = d_1 < d_2 < \cdots < d_{\tau(n)} = n$, the consecutive divisors of n. One of my oldest conjectures stated that almost all integers have two divisors satisfying

$$d_i < d_{i+1} < 2d_i . \tag{2}$$

For a long time (2) resisted all attempts. Recently Maier and Tenenbaum proved (2). In fact we have for almost all n

$$(\log n)^{1-\log 3-\varepsilon} < \min_i d_{i+1}/d_i < (\log n)^{1-\log 3+\varepsilon} . \tag{3}$$

The lower bound in (3) is due to R.R. Hall and myself, the upper to Maier and Tenenbaum. In fact we proved slightly better bounds than (3).

*) In the Erdös, Pomerance, Sárközy paper it will be shown that the number of $n \leq x$ with $\varphi(n) = \varphi(n+1)$ is at most $x/\exp(c(\log x)^{1/3})$.

Perhaps one will be able to get a distribution function for $\min_i d_{i+1}/d_i$. Perhaps the methods of Maier and Tenenbaum will show that for every ε and almost all n there are two consecutive divisors of n satisfying

$$d_i(1+\varepsilon) < d_{i+j} < d_i(1+2\varepsilon) \tag{4}$$

and that there is a sequence of consecutive divisors $d_i, d_{i+1}, \ldots, d_{i+t}$ satisfying

$$d_{i+j+1}/d_{i+1} < 1+\varepsilon \text{ for every } 1 \leqq j \leqq t \text{ and } d_{i+t} > 2d_i \ , \tag{4'}$$

i.e. the interval $(d_i, 2d_i)$ is covered by a mesh of close divisors. It is not clear to me at the moment how long a sequence of consecutive divisors one can get for which for every $1 \leqq j \leqq t_\varepsilon$

$$d_{i+j}(1+\varepsilon) > d_{i+j+1}$$

and in fact how does t_ε depend on ε ?

Denote by $\mu(n)$ the function of Möbius i.e. $\mu(n) = 0$ if n is divisible by a square and $\mu(n) = (-1)^{\omega(n)}$ if n is squarefree, where $\omega(n)$ is the number of distinct prime factors of n. Put

$$m(n) = \max_x \sum_{\substack{d|n \\ d<x}} \mu(d) \ .$$

Hall and I conjectured that for almost all n

$$m(n) \to \infty \ .$$

Perhaps Maier has a proof of this conjecture.

3. At a number theory meeting held at the University of Texas, Austin Texas in June 1982 I stated the following conjecture: There are infinitely many integers n for which

$$\sum_i \left(\frac{d_{i+1}}{d_i} - 1\right)^2 < C \tag{5}$$

where C is an absolute constant independent of n. A few weeks later this conjecture was proved by Vose. In fact he proved (5) with the exponent 2 replaced by $1+\varepsilon$, but then of course C must be replaced by C_ε. I further conjectured that (5) holds if $n = k!$ or if n is the product of the first k primes. This conjecture was proved by Tenenbaum by a modification of the method of Vose. Denote by $\tau(n)$ the number of divisors of n. I observed that if (5) holds for a sequence of integers $n_1 < n_2 < \cdots$ then in fact we must have

$$\tau(n_i)/(\log n_i)^2 \to \infty . \tag{6}$$

The proof of (6) requires only simple inequalities and can be left to the reader. I further asked: Let $g(n)$ tend to infinity arbitrarily slowly. Is there then a sequence n_i satisfying (5) and also

$$\tau(n_i) < g(n_i)(\log n_i)^2 \ ? \tag{7}$$

Vose proved a slightly weaker form of (7). He proved that there is a sequence satisfying (5) for which $\tau(n_i) < (\log n_i)^\alpha$ for some $\alpha > 2$.

Denote by $N(a, b)$ the least t for which

$$\frac{a}{b} = \frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_t}, \quad 1 < x_1 < \cdots < x_t$$

is possible. Put $N(b) = \max\limits_{1 \le a < b} N(a, b)$.

I proved

$$N(a, b) < c \frac{\log b}{\log\log b} .$$

In the proof I used that every $m < n!$ is the distinct sum of at most n divisors of $n!$. The results of Vose permit us to replace n by n^α for some $\alpha < 1$ and this gives

$$N(b) < (\log b)^c. \tag{8}$$

I conjectured $N(b) < c \log\log b$ which if true is best possible.

4. An old conjecture of mine states that $\binom{2n}{n}$ is never squarefree for $n > 4$. Sárközy recently proved that $\binom{2n}{n}$ can be squarefree only for a finite number of values of n and in fact he proves that the number of primes p for which $p^2 | \binom{2n}{n}$ tends to infinity with n. Probably his proof will show that for every α there is a prime p for which $p^{\alpha} || \binom{2n}{n}$ for all $n > n_0(\alpha)$.

Denote by $f(n)$ the number of integers k, $0 < k < n$ for which $\binom{n}{K}$ is squarefree. It is easy to see that $f(n) = 0$ for infinitely many n but that $f(n)$ can be arbitrarily large. I once conjectured that $f(n) = 0(n^{\varepsilon})$ for every $\varepsilon > 0$. As far as I know this conjecture is still open.

It is easy to see that the density of integers n, for which $f(n) = t$, exists and if we denote it by α_t, then $\sum_{t=0}^{\infty} \alpha_t = 1$.

About 10 years ago in a paper dedicated to D. H. Lehmer, Graham, Ruzsa, Straus and I proved among other things that for any two primes p and q there are infinitely many integers n for which $(\binom{2n}{n}, p\cdot q) = 1$. We could not prove this for $p\cdot qr$ and in particular we could not prove that there are infinitely many integers n for which $(\binom{2n}{n}, 105) = 1$. We conjectured that for some constant c

$$\sum_{\substack{p \nmid \binom{2n}{n} \\ p<n}} \frac{1}{p} < c \tag{9}$$

5. On p.17 of I the following conjecture is stated: "Is it true that for any partition of the pairs of positive integers into two classes the sums $\sum_{x \in X} \frac{1}{\log x}$ are unbounded where X ranges over all subsets which have all pairs belonging to one class?".

Rödl recently proved this conjecture and he further proved that this no longer holds for division into three classes.

II

Now I discuss some of my favorite old problems which are still unsolved.

1. Perphaps my favorite conjecture goes back to 1934. Is it true that for every integer c there is a finite system of congruences

$$a_i (\mathrm{mod}\ n_i) \quad c = n_1 < n_2 < \cdots < n_k \tag{10}$$

so that every integer satisfies at least one of the congruences (10)? I offer 1000 dollars for a proof or disproof of (10). This conjecture and some of its ramifications are extensively discussed in I and in my lecture given in Ooty; thus I will only mention questions not discussed in my recent papers.

Graham and I posed the following problem: Is it true that if one divides the integers greater than 1 into r classes then for at least one class one can find moduli which form the moduli of a covering system? This conjecture, if true, will no doubt be very difficult. We also conjectured that the equation

$$\Sigma \frac{1}{x_i} = 1 \quad x_1 < x_2 < \cdots < x_n \tag{11}$$

is solvable with all the x_i in the same class. This conjecture is probably much easier and has a much better chance of being true. Perhaps both conditions are satisfied in the same class. Perhaps these two conjectures have really nothing to do with dividing the integers into r classes but amount really to a density theorem. In fact perhaps the following result holds: For every $\varepsilon > 0$ there is an x_0 so that if $x \geq x_0$ and $1 < n_1 < n_2 < \cdots < n_t \leq x$ satisfy

$$\Sigma \frac{1}{n_i} > \varepsilon \log x$$

then one can select among the n_i the moduli of a covering system and also a solution of (11).

One can also formulate these problems as extremal problems: Determine or estimate the smallest $f(x)$ (respectively $g(x)$) so that if

$$\sum_{n_i<x} \frac{1}{n_i} > f(x) \tag{12}$$

then one can select among the n_i a set of solutions of (11) or if $f(x)$ is replaced by $g(x)$ then one can select the moduli of a covering congruence. I would expect that $g(x)$ is very much larger than $f(x)$. In any case the primes show that both must be larger than loglog x. I expect that both are very much larger. These problems perhaps change their character if we consider the set of all positive integers instead of the integers not exceeding x. Thus we state in I that perhaps any infinite sequence of positive upper density contains both the moduli of a covering system and solutions of (11). As far as I know these attractive and perhaps not hopeless questions have never been seriously considered.

In I we further state the following related problems: Is it true that if the integers are split into r classes then some class contains three distinct integers x, y, z satisfying $\frac{1}{x}+\frac{1}{y}=\frac{1}{z}$? Is it true that one class always contains (finitely many) integers whose reciprocals sum to each rational $\frac{a}{b}$? Observe that we have to restrict ourselves to finitely many integers since it is a simple exercise to show that if $x_n > 0$, $x_n \to \infty$ $\Sigma \frac{1}{x_n} = \infty$ then for every $\alpha > 0$ there is a subsequence for which $\Sigma \frac{1}{x_{n_i}} = \alpha$. As far as I know the following problems have not yet been considered. Determine or estimate the smallest $F(x)$ (respectively

$f(x))$ for which if

$$A(x) = \sum_{a_i<x} 1 > F(x) \left(\text{resp. } \sum_{a_i<x} \frac{1}{a_i} > f(x)\right)$$

then the sums $\frac{1}{a_i} + \frac{1}{a_j}$ cannot all be distinct. The primes show that $F(x) > c\,\frac{x}{\log x}$, $f(x) >$ loglog x, but I do not know the true order of magnitude of these functions.

A system of congruences is called disjoint if no integer satisfies more than one of them. Stein and I conjectured that if $a_i \pmod{n_i}$ with $n_1 < n_2 < \cdots < n_r \le x$ is a disjoint system of congruences then $r = 0(x)$. This was proved by Szemerédi and myself. We in fact showed that if $f(x) =$ max r then for some $c > 0$

$$\frac{x}{\exp(\log x)^{1/2 +\varepsilon}} < f(x) < \frac{x}{(\log x)^c}. \tag{13}$$

We conjectured that the lower bound in (13) is closer to the truth but we could get nowhere with this conjecture.

I conjectured, and Mirsky and Newman proved, that there is no disjoint system of covering congruences. Here I would like to mention the following nice and much more general problem of Herzog and Schönheim. If G is an abelian group, can there be an exact covering of G by cosets of different sizes?

Let $a_i \pmod{n_i}$ be a disjoint set of congruences. Assume that all the n_i are larger than t. Can one estimate max $\sum_{n_i>t} \frac{1}{n_i} = h(t)$? Is it true that $h(t) \to 0$ as $t \to \infty$? Perhaps I overlook a trivial argument.

Let $\tau'(n)$ denote the cardinality of the largest set of disjoint congruences formed from the divisors of n. Can one get a non-trivial estimate for $\tau'(n)$ and can one get an asymptotic formula for $\sum_{n=1}^{x} \tau'(n)$?

For every integer n there is a real c_n defined as follows: From all divisors d_i of n form all possible disjoint systems $a_i \pmod{d_i}$. Let c_n be the greatest lower bound of the densities of the set of integers not satisfying any of these congruences. $c_n > 0$ is clearly a rational number. It is not difficult (but not trivial) to prove that the sequence c_n has a distribution function but I cannot say very much about it and in particular I cannot estimate c_n in terms of the divisors of n.

In I we define a number n to be covering if one can form a covering system among the divisors of n that exceed 1. The covering numbers have a density. Benkoski and I asked: Is there a C so that if $\sigma(n)/n > C$ then n is covering? J. Haight showed that no such C exists. His ingenious proof is quite difficult.

Benkoski and I conjectured that if C is sufficiently large and if $\sigma(n)/n > C$ then n can be written as the sum of proper divisors of n. This conjecture is still open.

Very recently the following question occurred to me. Is it true that for every r there are integers n_r so that if we divide the divisors of n_r in any way into r classes then n_r always is the distinct sum of divisors of the same class? If the answer is affirmative, is there a C_r so that every n with $\sigma(n)/n > C_r$ has this property? I cannot answer these questions even for $r = 2$.

2. Now I discuss some problems connected with van der Waerden's theorem. This is also discussed in I and in the excellent book of Graham, Rothschild, and Spencer. Thus I will try to keep my report short. Let $f(n)$ be the smallest integer for which if one divides the integers not exceeding $f(n)$ into two classes, at least one class contains an arithmetic progression of n terms. It is a surprising fact that $f(n)$

is very hard to estimate both from above and below. The upper bound given by van der Waerden's proof gives that $f(n)$ increases not faster than Ackerman's well known function which is not primitive recursive. The best lower bound is $f(p) > p\, 2^p$ (p prime) due to Berlekamp and $f(n) > c2^n$ due to Lován and myself. As far as I know $f(n)/2^n \to \infty$ has never been proved. I offered 100 dollars for a proof of

$$(14) \qquad f(n)^{1/n} \to \infty .$$

I now raise this offer to 250 dollars. I will give 100 for $f(n) > (2+\varepsilon)^n$ and 25 for $f(n)/2^n \to \infty$. It is surprising that the proof of (14) seems to present serious difficulties. I am convinced that (14) is true and will pay 10000 dollars for a disproof.

As far as I know, the first person to suggest that perhaps $f(n)$ really increases as fast as Ackermann's function was Solloway in a discussion with Graham a few years ago. At first I thought that this is complete nonsense but in view of the surprising results of Paris-Harrington I realized that Solloway's suggestion should be taken very seriously. In any case I give 100 dollars for a proof that $f(n)$ is primitive recursive and 500 dollars for a proof that it is not.

About 50 years ago Turán and I guessed that van der Waerden's theorem is really a density theorem and has nothing to do with the division of the integers into classes. In fact denote by $r_k(n)$ the smallest integer for which every sequence $1 \leq a_1 < a_2 < \cdots < a_\ell \leq n$, $\ell = r_k(n)$ contains an arithmetic progression of k terms. We conjectured

$$(15) \qquad \lim r_k(n)/n = 0.$$

The bound $r_k(n) = \frac{n}{2}$ for $n > n_0(k)$ would clearly imply van der Waerden's theorem. We hoped that a good estimation of $r_k(n)$ would permit us to get a good bound for $f(n)$. We at first did not realize that (15) is a very hard problem. The first indication that (15) was difficult came up when Salem and Spencer showed

$$r_3(n) > n^{1-c/\log\log n} .$$

We at first thought that $r_k(n) = n^{1-c_k}$. The best current bounds for $r_3(n)$ are

$$n \exp(-c(\log n)^{1/2}) < r_3(n) < cn/\log\log n. \tag{16}$$

(16) has not been improved for more than 30 years.

The upper bound in (16) is due to K. F. Roth and the lower to F. Behrend. I offered 1000 dollars for (15) and late in 1972 Szemerédi found a brilliant but very difficult proof of (15). I feel that never was a 1000 dollars more deserved. In fact several colleagues remarked that my offer violated the minimum wage act. Later Furstenberg obtained a proof of (15) by using ergodic theory. There is no doubt that ergodic theory will be useful to settle many further problems in combinatorial number theory and perhaps it will eventually rival the applications of algebra and analysis to number theory. It would be very desirable to obtain as good upper and lower bounds for $r_k(n)$ as possible. To ask for an asymptotic formula may be, to quote P. Elliott's excellent book, Probabilistic Number Theory, like "baying at the moon". The question appears clearly hopeless and beyond human intelligence at the moment. Unfortunately our original hopes with Turán that the study of $r_k(n)$ will give an estimation for $f(n)$ have so far not been fulfilled. Szemerédi's proof uses van der Waerden's theorem and Furstenberg's proof is a pure existence proof.

An old conjecture in number theory states that for every k there are k primes in an arithmetic progression. Prichard recently found 18 primes in an arithmetic progression, which is the current record. I conjectured about 40 years ago that if

$$\Sigma \frac{1}{a_k} = \infty \tag{17}$$

then the a's contain arbitrarily long arithmetic progressions. I offer 3000 dollars for a proof or disproof of this conjecture. (17) at the moment seems unattackable but if "we" live, it will perhaps be settled in the next century ("We" here stands for humanity). It is not even known that if (17) holds then $a_i + a_j = 2a_r$ is solvable among the a's i.e. the a's contain an arithmetic progression of three terms.

It is easy to see that if my conjecture holds then there is a $g(k)$ so that if

$$\Sigma \frac{1}{a_i} > g(k) \tag{18}$$

then the a's contain an arithmetic progression of length k. It would be very interesting to estimate $g(k)$ as accurately as possible. Trivially

$$g(k) > \frac{1}{2} \log f(k). \tag{19}$$

It would of course be very interesting if one could get an upper bound for $g(k)$ in terms of $f(k)$ or if one could improve the lower bound (19). Graham and I could not even prove $g(k) > (\frac{1}{2} + \varepsilon) \log f(k)$. In fact it seemed quite possible to us that

$$\lim g(k)/\log f(k) = \infty .$$

Gerver proved that

$$g(k) \geq (1+0(1))k \log k \tag{20}$$

and he thought that perhaps (20) could be best possible. If true this certainly would be a sensational result.

As far as I know the following question has never been investigated: Assume that (18) holds. Denote by $g(n; k)$ the upper bound $\Sigma \frac{1}{a_i}$ where $n < a_1 < a_2 < \cdots$ and the a's do not contain any arithmetic progression of k terms. It would be very nice if we could estimate

$$\lim_{n \to \infty} g(n; k). \tag{21}$$

Perphas the limit (21) is 0 for every k. I have no idea how to attack this question, but perhaps I overlook a simple argument.

Szemerédi observed that we cannot even prove that

$$r_4(n)/r_3(n) \to \infty.$$

Perhaps even the proof of $r_4(n) - r_3(n) \to \infty$ is not completely trivial.

Just like the Ramsey numbers we can define van der Waerden numbers. $r(u, v)$ is the smallest integer so that if we split the integers not exceeding $r(u, v)$ into two classes either the first class contains an arithmetic progression of u terms or the second class contains an arithmetic progression of v terms. This led me to the following question. Divide the integers not exceeding n into two classes $s_1(n)$ and $S_2(n)$. Assume $|S_1(n)| = k$. Denote by $h(n; k)$ the largest integer so that $S_2(n)$ must contain an arithmetic progression of $h(n; k)$ terms. Trivially $h(n; k) \geq \frac{n}{k+1}$ and Hegyván observed that for small k $(k < \frac{c \log n}{\log\log n})$ this is best possible. The most interesting case seems to be $k = [n^{1/2}]$. I hope that

$$h(n; n^{1/2})/n^{1/2} \to \infty . \tag{22}$$

I made no progress with (22) but perhaps I again overlook a simple idea.

Finally the following problem of Graham and myself is of interest. Estimate the length of the longest arithmetic progression

$$a + kd,\ 0 \leq k \leq t,\ a + td \leq x \tag{23}$$

which consists entirely of primes. It is easy to see that $t < (1 + O(1))\log x$ and in I we conjecture that $t = O(\log x)$ but we could not even prove $t < (1-\varepsilon)\log x$.

Let us slightly modify the problem. We only insist that $d \leq x$. Then perhaps there always is an a so that there is an arithmetic progression (23) of length $(1 + O(1))\log x$ all of whose terms are primes, or more generally let p be the smallest prime that does not divide d. I expect that there always is an a so that there is a progression (23) of length $p-1$ (that does not contain p) all of whose terms are primes. If true this conjecture is clearly best possible but it is also clear that there is no hope in the foreseeable future to prove or disprove it. Pomerance noted that this follows from the prime k-tuples conjecture.

We have much more success if one asks for a progression which consists entirely of squarefree numbers. In fact we prove the following

<u>Theorem</u>. Let d be an integer and p the smallest prime with $p \nmid d$, then there is a progression

$$a + kd,\quad 0 \leq k < p^2 - 1 \tag{24}$$

of $p^2 - 1$ integers all of which are squarefree.

(24) is clearly best possible. It will be clear from the proof that a can be chosen in any residue class mod d satisfying $(a, d) = 1$. To prove our theorem, assume that $p = p_\ell$, i.e. $p_\ell \nmid d$ but $p_j | d$ for all $j < \ell$; clearly if $(a, d) = 1$ then $a + ud \not\equiv 0 \mod p_j$ for $j < \ell$. Let now y be large compared with d and let x be large compared with y.

First of all observe that there is a constant $\alpha > 0$ so that the number of integers $a < x$, $(a, d) = 1$ for which all the integers

$$a + kd,\quad 0 \leq k < p_\ell^2 - 1 \tag{25}$$

are not divisible by p_j^2, $\ell < j < y$, is greater than αx where α is independent of y and x. To see this, first of all note that we already have that none of the integers (25) are divisible by any prime $p < p_\ell$. A simple sieve process now gives that

$$\vartheta > x \prod_{i=1}^{\ell-1} (1 - \frac{1}{p_i}) \prod_{j=\ell}^{y} \left(\frac{p_\ell^2 - 1}{p_j^2} \right) > \alpha x \ . \tag{26}$$

Now the number of integers $a < x$ for which one of the integers (25) are divisible by the square of a prime $p > p_t$ is clearly less than

$$x \sum_{j>t} \frac{p_\ell^2 - 1}{p_j^2} < \frac{x}{p_t} \ . \tag{27}$$

From (26) and (27) we obtain that the number of integers $a < x$ for which all the integers (25) are squarefree is greater than

$$(\alpha - \frac{1}{p_t})x > 0 \tag{28}$$

if p_t is sufficiently large. Thus our theorem is proved. Clearly an analogous theorem holds if p_k^2 is replaced by any sequence of primes the sum of whose reciprocals converges.

Our problem becomes much less trivial if we ask for the longest progression (23) all terms of which are squarefree (i.e. we again insist that $a + td \leqq x$). Here we can only show that there is a progression (23) of length $c\log x$ all terms of which are squarefree and we cannot decide if this is best possible. Perhaps $\log x$ can be replaced by $(\log x)^2$. It is easy to see that there is no such progression of length $(1+\varepsilon)(\log x)^2$.

3. Now I discuss some problems on consecutive integers. I wrote a great deal about this subject and here I only mention a few problems which I like and which have perhaps been neglected. Denote by $P(m)$ the greatest

and by $p(m)$ the least prime factor of m. Put

$$n + i = a_i b_i \quad \text{where} \quad P(a_i) \leq k \quad \text{and} \quad p(b_i) > k \ .$$

Denote by $f(n;\ k)$ the largest index ℓ for which all the integers $a_1, \ldots, a_\ell$ are distinct. Gordon and I proved that

$$(29) \qquad \limsup_{n,k} f(n;\ k)/k \leq 2.$$

In other words there are only a finite number of integers n and k for which $f(n;\ k) > (2+\varepsilon)k$. Probably in (29) 2 can be replaced by 1 but we never could improve (29). Denote by p_r the least prime greater than k. Then perhaps

$$(30) \qquad f(k) = \max_n f(n;\ k) = p_{r+1} - 1.$$

The integers $2, 3, \ldots, p_{r+1}-1$ show that $f(k) \geq p_{r+1}-2$ and perhaps (30) holds for all k (perhaps with a finite number of exceptions).

I conjectured that there is an absolute constant $c > 0$ so that for every n and k the number of distinct integers in the sequence $a_1, a_2, \ldots, a_k$ is greater than ck.

I further conjectured that for every $\varepsilon > 0$ there are only finitely many integers n and k for which

$$(32) \qquad \min_{1 \leq i \leq \ell} a_i > \varepsilon k.$$

Ruzsa observed that if (32) is true, it is nearly best possible since there are infinitely many integers n and k for which

$$f(n;\ k) < ck/\log k.$$

These two conjectures seem attractive and not hopeless and I hope they will be settled in a finite time.

Finally I mention two of my problems. One is one of my favorite old problems. Denote by

$$1 = a_1 < a_2 < \cdots < a_{\varphi(n)} = n-1$$

the integers relatively prime to n. Is it true that there is an absolute constant c so that

$$(33) \qquad \sum_{i=1}^{\varphi(n)-1} (a_{i+1} - a_i)^2 < c \, \frac{n^2}{\varphi(n)} \; ?$$

Hooley has some interesting partial results (see I). It is surprising that (33) resisted so far all attacks and I offer 500 dollars for a proof or disproof of (33).

To state the other problem denote by

$$J(n) = \max(a_{i+1} - a_i)$$

$J(n)$ is named after Jacobstahl who started the investigation of $J(n)$. Let x be large. Is it true that there are two integers a and $b \vartheta$ satisfying

$$(34) \qquad a < b < x, \ (a, b) = 1, \ J(a) > \log x, \ J(b) > \log x \ ?$$

It is easy to see that in (34) the answer is affirmative if $\log x$ is replaced by $\frac{\log x}{\log\log x}$. Perhaps (34) should be reformulated as follows: Put

$$(35) \qquad h(x) = \max \min(J(a), J(b))$$

where in (35) the maximum is to be taken over all pairs $(a, b) = 1$, $a < x$, $b < x$. It is an easy exercise of the application of the sieve of Eratosthenes to show that

$$h(x) > c \log x/(\log\log x)^{1/2} \; .$$

At the moment I cannot improve this and what is perhaps worse I have no non-trivial upper bound for $h(x)$.

4. Now I discuss some problems on prime numbers. I will of course not discuss the classical problems except the lower bound of $p_{k+1} - p_k = d_k$. The best known lower bound for d_k is due to R.A. Rankin. He proved in 1938 that there is a $c > 0$ so that for infinitely many k

$$(36)\qquad p_{k+1}-p_k = d_k > \frac{c \log k \operatorname{loglog} k \operatorname{loglogloglog} k}{(\operatorname{logloglog} k)^2} = L_k .$$

4 years earlier I proved (36) without the last factor logloglog log k. (36) seems to be the natural boundary of the method of Rankin and myself. The only improvement of (36) during the last 45 years is that Schönhage and later Rankin himself improved the value of c. This fact made me offer 10000 dollars for a proof that (36) holds for every c. (For a disproof of this conjecture, which is of course out of the question I offer 25000 dollars.) I could just as well have offered 10 dollars since the conjecture is surely true, but in the unlikely case that I am wrong I want to be able to pay what I promised and I believe I could earn, beg, borrow or steal 25000 dollars. Perhaps the following little story will show our powerlessness with the problems about primes. A little more than 60 years ago I learned from my father the proof of Euclid that there are infinitely many primes and that there are arbitrarily large gaps between the primes since $n!+2,\dots,n!+n$ are all composite. Later I realized that this simple idea which really every baby should understand gives $d_k > \frac{\log k}{\operatorname{loglog} k}$ which is not very much weaker than (36).

In 1939 Cramér conjectured that

$$(37)\qquad \limsup d_k/(\log k)^2 = 1.$$

Cramér explains that he arrived at (37) by observing that if we define a random sequence where the probability that n belongs to the sequence is $\frac{1}{\log n}$, then it easily follows from well known results of probability theory that (37) holds for almost all such sequences. Needless to say, by these arguments one can never probe (37). I asked Maier whether he can disprove the following (very unlikely) conjecture:

$$(38)\qquad (\pi(x+t) - \pi(x))/(t/\log x) \to 1$$

if $x \to \infty$ and $t > (\log x)^{1-\varepsilon}$ for some $\varepsilon > 0$. (38) of course contradicts (37) but is compatible with (36). In fact Maier did much more. He proved that for every $\ell > 1$ there is an $\varepsilon = \varepsilon_\ell$ and two sequences $x_n \to \infty$, $y_n \to \infty$ so that

$$(39)\qquad \lim(\pi(x_n+(\log x_n)^\ell)-\pi(x_n))/(\log x_n)^{\ell-1} > 1+\varepsilon$$

and

$$(40)\qquad \lim(\pi(y_n+(\log y_n)^\ell)-\pi(y_n))/(\log y_n)^{\ell-1} < 1-\varepsilon\ .$$

Thus the primes are not as uniformly distributed as one could have guessed and it seems to me that (39) and (40) contradict the heuristic assumption that the distribution of primes in the large can be described best by studying the random sequences when n is in the sequence with probability $\frac{1}{\log n}$. Perhaps if $y \to \infty$ and z tends to infinity faster than any power of log y then uniformly in y and z

$$(41)\qquad (\pi(y+z) - \pi(z))/\frac{z}{\log y} \to 1\ .$$

There is of course no hope to prove (41), but if it is false perhaps one could disprove it.

I conjectured that for every r there is a $c_r > 0$ so that for infinitely many values of k

$$(42)\qquad \min_{1\le i\le r} (p_{k+i+1} - p_{k+i}) > c_r L_k\ .$$

I proved (42) for $r = 2$ but could not do it for $r > 2$. Maier in a recent paper found an ingenious proof of my conjecture in the general case. In his proof $c_r \to 0$ as $r \to \infty$, but in fact probably (42) holds for every c. Put

$$\max_{p_n<x}\ \min_{1\le i\le r}\ (p_{n+i+1} - p_{n+i}) = f(n,\ r).$$

I would expect that

$$(43) \qquad f(n, r)/f(n, 1) \to 0$$

as $n \to \infty$. (43) is probably hopeless for the present even if the primes are replaced by squarefree numbers. I further conjectured that for every r

$$(44) \qquad \liminf \max(d_{k+1}, d_{k+2}, \ldots, d_{k+r})/\log k < 1$$

but could prove (44) only when $r = 1$ and for $r > 1$ (44) is still open.

Turán and I easily proved that $d_k > d_{k+1}$ and $d_{k+1} > d_k$ both have infinitely many solutions. We noticed that we cannot prove that $d_{k+2} > d_{k+1} > d_k$ also has infinitely many solutions. In fact we could not even prove that for $k > k_0$, $(-1)^r(d_{k+r+1} - d_{k+r})$ cannot always have the same sign, perhaps we overlook a trivial idea. I offer 100 dollars for a proof that this is impossible and 100000 dollars for a proof that our conjecture is wrong (which is of course completely out of the question). Presumably $d_k = d_{k+1}$ has infinitely many solutions, but this is probably hopeless at the present state of science.

A few days ago Pomerance and I started to investigate the following question: Consider the r integers $d_{k+1}, d_{k+2}, \ldots, d_{k+r}$ and order them by size. This gives a permutation of the integers $1, 2, \ldots, r$. One would of course expect that all the $r!$ permutations will occur, but in view of the difficulties we had with Turán, a proof of this conjecture cannot be expected. Denote by $f(r)$ the number of permutations which occur infinitely often. So far we could not prove anything better than that $f(r) \geqq r$ (which is very simple).

By the way it is not difficult to prove that if the primes are replaced by squarefree integers then all the $r!$ permutations must occur.

Denote by $P(n)$ the greatest prime factor of n. Pomerance and I proved that $P(n) < P(n+1) < P(n+2)$ has infinitely many solutions and here in fact we can prove for $r = 3$ that there are at least $r+1$ permutations which must occur infinitely often and perhaps our proof will work for $r > 3$.

Incidentally it is not difficult to prove that with $\varphi(n)$, $d(n)$, $\sigma(n)$ all the $r!$ permutations occur and in fact all occur with positive density.

Perhaps the following question which just occurs to me might be of interest. To fix our ideas let us consider $\omega(n)$, the number of distinct prime factors of n. If $h(n)$ tends to infinity sufficiently slowly then for almost all n all the integers $\omega(n+i)$, $1 \leq i \leq h(n)$ will be different. If $h(n)$ tends to infinity fast enough then for almost all n the integers $\omega(n+i)$, $1 \leq i \leq h(n)$ will not be all different. Can we find $h(n)$ so that the density of the integers n for which the integers $\omega(n+i)$ are all different is $\frac{1}{2}$ (or, more generally is c) and how does then c determine $h(n)$? The same question can be asked for the other functions too, but probably we have most hope for answering this question for $\omega(n)$ and there is very little hope for a reasonable answer in case of, say, $\varphi(n)$.

To end this chapter let me retract an old statement of mine. Let α be an irrational number. No doubt the sequence $p_n\alpha - [p_n\alpha]$ is not well distributed and in fact there is no doubt that for every irrational α and $h > h_0(\varepsilon)$ and $\varepsilon > \alpha$ there is an n so that for all $0 < i < k$

$$0 < p_{n+i}\alpha - [p_{n+i}\alpha] < \varepsilon. \tag{44}$$

(44) is clearly unattackable at present. I claimed that I can prove that there is an irrational α for which the sequence $p_n\alpha - [p_n\alpha]$ is not

well distributed. The theorem is no doubt correct and perhaps will not be difficult to prove but I never was able to reconstruct my "proof" which perhaps never existed.

The notion of a well distributed sequence is due to Hlawka (for more information see the nice book of Kuipers and Niederreiter).

Let $0 < x_n < 1$. The sequence $\{x_n\}$ is well distributed if to every $\varepsilon > 0$ there is a k so that for every n and $0 < a-b < 1$ the number of indices i for which $a < x_{n+i} < b$, $1 \leq i \leq k$ is between $k(b-a-\varepsilon)$ and $k(b-a+\varepsilon)$.