SOME PROBLEMS ON NUMBER THEORY

P. Erdös

This lecture was given at a meeting on number theory held at Marseille (Luminy). I added a few new problems and also of course added the new results which were obtained in the meantime.

In this little note I discuss mainly problems on prime numbers. Some of these have occupied me for a long time, but I mention also some new questions. The quality of the problems considered will be very uneven; some are more exercises, some certainly serious problems. Unfortunately, I am not always sure into which category the problems belong.

First I discuss some problems which arose during our meeting in Marseille. An old conjecture of Mirsky and myself [1] states that ($d(n)$ denotes the number of divisors of n) $d(n) = d(n+1)$ has infinitely many solutions. It is probably presumptious to call it "our conjecture"; it probably was asked long ago. I only call it our conjecture since it is mentioned in one of our papers. Brun's method gives that for infinitely many n, $c_1 < d(n)/d(n+1) < c_2$, and in fact the set of limit points of $d(n)/d(n+1)$ contains intervals [2]. No doubt the sequence $d(n)/d(n+1)$ is everywhere dense in $(0,\infty)$, but the only limit points known at that time were 0 and ∞. Our original conjecture on $d(n) = d(n+1)$ seemed to be unattackable, and it was a great surprise to me when Claudia Spiro (unpublished) proved that $d(n) = d(n+5040)$ has infinitely many solutions. Recently, by using and further developing the method of Claudia Spiro, Heath-Brown [3] proved that $d(n) = d(n+1)$ has infinitely many solutions. In fact he proved that the number of solutions of $d(n) = d(n+1)$, $n < x$ is $> cx/(\log x)^7$. Pomerance, Sárközy and I proved that the number of solutions is $< cx/(\log\log x)^{\frac{1}{2}}$ [4]. No doubt this gives the correct order of magnitude. As far as I know, the problem of whether the set of limit points of $d(n)/d(n+1)$ is everywhere dense in $(0,\infty)$ is still open.

The proof of Claudia Spiro is based on the fact that there are 8 primes p_i, $i = 1,\ldots,8$ so that the least common multiple of the differences $p_j - p_i$, $1 \le i < j \le 8$ is 5040. This led Narkiewicz and me to consider the following problem: Denote by $D(p_1,\ldots,p_n)$ the least common

CONGRESSUS NUMERANTIUM **54**(1986), pp.225-244

multiple of the $\binom{n}{2}$ numbers $p_j - p_i$. Put

$$f(n) = \min_{p_1,\ldots,p_n} D(p_1,\ldots,p_n)$$

and let F(n) be the smallest value of $D(p_1,\ldots,p_n)$ assumed for infinitely many $p_1,p_2,\ldots,p_n$. We of course can not even prove that F(2) is finite, since this would imply that $p_{k+1} - p_k < C$ has infinitely many solutions for some C, but we will assume the prime k-tuple conjecture of Hardy and Littlewood, which of course implies $F(n) < \infty$. Put

$$g(n) = \prod_{q<n} q^{\alpha_q}$$

where α_q is the largest integer for which $\phi(q^{\alpha_q}) = (q-1)q^{\alpha_q-1} < n$. A simple argument shows that $F(n) \geq g(n)$, since if q is not one of the p's, then $q^{\alpha_q} | D(p_1,\ldots,p_n)$. If q is one of the p's then $q^{\alpha_q} | D(p_1,\ldots,p_n)$ if $(q-1)\cdot q^{\alpha_q-1} \leq n-1$. We conjectured that $f(n)/g(n) \to \infty$ and that $f(n) = g(n)$ is possible only for very small values of n. Perhaps $f(n) = F(n)$ for $n > n_0$. We could not even show that $f(8) = 5040$. It could be 2520 if all the 8 p's are incongruent mod 16. We only could exclude this by long computations which we did not carry out. It follows from the prime number theorem that $\log g(n) = (1+o(1))n$. We think that perhaps

$$(1) \qquad \lim \frac{\log f(n)}{n} < \infty, \quad \lim \frac{\log F(n)}{n} < \infty .$$

It might be of some interest to obtain an asymptotic formula for $\log D(2,3,\ldots,p_n)$; probably,

$$(2) \qquad \log D(2,3,\ldots,p_n)/n \log n = c, \text{ for some } 0 < c < 1.$$

In a recent letter Claudia Spiro deduced from the prime k-tuple conjecture that

$$(3) \qquad F(n) < (g(n))^{1+c\frac{\log\log n}{\log n}} .$$

(3) of course implies (1). The conjecture $F(n)/g(n) \to \infty$ and $f(n)/g(n) \to \infty$ remains open. In view of her result (3) it would perhaps

be of interest to study

$$\min_{p_1,\ldots,p_n} \{(\max_{1<i<n} p_i) D(p_1,\ldots,p_n)\} = A_n.$$

It is true that $A_n^{1/n} \to \infty$? or at least $A_n > (e+\varepsilon)^n$; i.e., $A_n > (\prod_{p_1<n} p_i)^{1+\varepsilon}$? A related function is

$$\min_{p_1,\ldots,p_n} D(p_1,\ldots,p_n) \prod_{i=1}^{n} p_i = B_n.$$

$B_n > (n!)^{1+\varepsilon}$ would perhaps be of some interest.

These problems can be considered for other sequences than the primes. If $a_1,a_2,\ldots,a_n$ are n square-free numbers what can be said about $\min D(a_1,\ldots,a_n)$? At the moment I can say nothing non-trivial about this problem.

Some questions which Nicolas and I considered lead to the following question: let $p_1,p_2,\ldots,p_n$ be an arbitrary set of n primes. Is it true that

$$\sum_{1<i<j<k} \frac{1}{p_j-p_i} < Cn \; ? \tag{4}$$

(4) is still open. It follows from the prime k-tuple conjecture that (4) if true is best possible; i.e., there are infinitely many n-tuples of primes $p_{i_1},\ldots p_{i_k}$ for which

$$\sum_{1<j<j'<n} \frac{1}{p_{i_j}-p_{i_{j'}}} > C_1 n \; .$$

I thought for a while that instead of (4) the following stronger result may hold: Let $a_1 < a_2 < \ldots < a_n$ be a sequence of integers for which every interval of length t contains for every t, fewer than $c_1 t/\log t$ a's. Is it then true that

$$\sum_{1<i<j<n} \frac{1}{a_j-a_i} < Cn \; ? \tag{5}$$

Unfortunately, Ruzsa gave a simple counterexample to (5). Let the a's be the integers of the form $\sum_{i=1}^{s} \varepsilon_i 2^i$, where $\varepsilon_i = 0$, or 1 but $\varepsilon_i = 0$ if i is a power of 2, and s is chosen so that $s - \frac{\log s}{\log 2} = \frac{\log n}{\log 2} + O(1)$.

It is easy to see that the a's satisfy our condition but

$$\sum_{1<i<j<n} \frac{1}{a_j - a_i} > c\, n \log\log n. \tag{6}$$

(6) contradicts (5) and is easily seen to be the best possible. Probably a counterexample to (4) can also be found (i.e., the a's can be chosen to be primes). Put $d_k = p_{k+1} - p_k$; d_k seems to behave very irregularly. Put

$$D(x) = \max_{p_k < x} (p_{k+1} - p_k) .$$

Cramer [5] conjectured that $\overline{\lim_k} \frac{d_k}{(\log k)^2} = 1$. A slight strengthening of Cramer's conjecture states

$$\lim \frac{D(x)}{(\log x)^2} = 1. \tag{7}$$

It is quite possible, though, that Cramer's conjecture holds but (7) is false. (7) in particular would imply that

$$\frac{D(2x)}{D(x)} \to 1$$

and there certainly is no real evidence that this holds. There is no doubt that every even d is of the form $p_{k+1} - p_k$, but the smallest for which $p_{k+1} - p_k = d$ probably tends to infinity exponentially in d, but it seems to be hopeless to prove that it tends to infinity faster than polynomially.

I now add some new conjectures: Denote by $D_1 < D_2 < \ldots$ the values of $\max_{p_k<x} (p_{k+1}-p_k)$ as $x \to \infty$. Perhaps the following problems are of some interest. 2,4,6,8,14... are the first few values of the D's. It seems certain that the density of the D's is 0, and perhaps $D_{k+1} - D_k \to \infty$, but on the other hand, perhaps $D_{k+1} - D_k = 2$ has infinitely many solutions. Also I expect that

$$D_{k+1}/D_k \to 1.$$

Let r_k be the smallest index for which

$$p_{r_k+1} - p_{r_k} = D_k .$$

I am sure that $r_{k+1} - r_k \to \infty$. In other words the abnormally large differences between consecutive primes are far apart. I should stop stating hopeless conjectures; quoting Hardy, "any fool can ask problems *on primes which no wise man can answer.*"

Denote by $U(x)$, the number of even integers of the form $p_j - p_i$, $3 < p_i < p_j < x$. $U(x) > cx$ follows immediately by Brun's method, but perhaps, $U(x) > \frac{x}{2} - (\log x)^{\alpha}$, for some α and all $x > x_0(\alpha)$, and perhaps for infinitely many x : $U(x) > \frac{x}{2} - C$ for some absolute constant C. Both of these conjectures are of course unattackable in the foreseeable future (the second one can perhaps be disproved), but I believe that if 1 is counted as a prime then there are infinitely many primes $\leq p$ so that every even number less than p can be written as the difference of 2 primes $\leq p$ i.e., $U(p) = \frac{p-1}{2}$; $U(p) = \frac{p-3}{2}$ if 1 is not counted as a prime.

Denote by $V(x)$ the number of integers of the form $a_j - a_i$ where $1 < a_i < a_j < x$ are squarefree numbers. $V(x) > x - x^{\alpha}$ is easy to prove for some $\alpha < 1$, also $V(x) > x - C$ holds for infinitely many x and it seems to be easy to prove that for every t, the density of the integers x for which $V(x) = x - t$ exists, and the density of integers x for which $V(x) < x - t$ tends to 0 as $t \to \infty$. The reason for the vagueness of my statement is that I did not think the proof over in all details. Rankin [6] proved in 1938 that $(\max_{p_k < x} (p_{k+1} - p_k) = D(x))$

$$(8) \qquad D(x) > c \log x \log\log x \log\log\log\log x (\log\log\log x)^{-2} = L(x).$$

Since then the only improvement of (8) was that the original value of c has been replaced by a larger one by Schönhage and Rankin. This fact lead me to offer a reward of 10^4 dollars for a proof that (8) holds for every c and infinitely many x (in fact it no doubt holds for all $x > x_0(c)$). I am so sure that this conjecture is true that I offer $25,000 for a disproof. I really feel like offering 10^6 dollars, but contrary to rumors [7], I never offer a prize if I could not pay it (and perhaps if necessary I could earn, beg, borrow or steal $25,000 dollars).

Let $H(x)/D(x) \to \infty$. Is it true that ($\pi(y)$ is the number of primes not exceeding y)

$$(9) \qquad \pi(x+H(x)) - \pi(x) > C' H(x)/\log x \ ?$$

(9) if true, is no doubt unattackable at present. Maier [8] recently proved that if

$$\pi(x+H_1(x)) - \pi(x) = (1+o(1))H_1(x)/\log x \tag{10}$$

then $H_1(x)/(\log x)^c \to \infty$ for every c. His ingenious proof is surprisingly simple. Perhaps if $H_1(x)/(\log x)^c \to \infty$ for every c then (10) holds, but this of course is again unattackable. Denote by A(x) the number of distinct integers of the form $p_{k+1} - p_k$, $p_k < x$. Is it true that

$$A(x)/D(x) \to 0 \ ? \tag{11}$$

I have no intuition about (11) and it is quite possible that the limit in (11) does not exist. I expect that

$$\max_{p_k<x} \min(p_{k+1}-p_k, p_k-p_{k-1}) / \max_{p_k<x} (p_{k+1}-p_k) \to 0 \ . \tag{12}$$

(12) is certainly true, but is probably very deep. All these questions can be formulated for the sequence $q_1 < q_2 < \ldots$ of squarefree numbers. Unfortunately these questions seem to me nearly as difficult as the questions about primes, with a few exceptions. It is a simple exercise in the use of the sieve of Eratosthenes, that for every d there are infinitely many indices k for which $q_{k+1} - q_k = d$. The smallest such probably increases exponentially in d: we can at least show that it does not increase faster. Let $p_1 < p_2 < \ldots$ be an infinite sequence of primes, $a_1 < a_2 < \ldots$ is the sequence of integers not divisible by any of the p's. We can ask the same questions about $a_{i+1} - a_i$ but can answer them only if the p's tend to infinity very fast.

Perhaps we have more chance for success if we consider the integers relatively prime to n. Let $1 = a_1 < \ldots < a_{\phi(n)} = n - 1$ be the integers relatively prime to n, and put (J(n) after Jacobstahl) [9]:

$$J(n) = \max_{a_i<n} (a_{i+1} - a_i) \ .$$

Jacobstahl conjectured $J(n) < c(\log n)^2$, and this was proved by Iwaniec [10], but perhaps $J(n) < (\log n)^{1+\varepsilon}$. This would require very much better sieve methods than the ones at our disposal at present.

Let n_k be the product of the first k primes. Jacobstahl conjectured that for $m \le n_k$, $J(m) < J(n_k)$. Perhaps $J(m) \le J(n_k)$ for all m n_{k+1}, with possibly a finite number of exceptions. Clearly $J(n_{k+1}) > J(n_k)$ and probably

$$J(n_{k+1}) - J(n_k) \to \infty \text{ but } J(n_{k+1})/J(n_k) \to 1. \tag{13}$$

The second conjecture of (13) seems certain to be true. The following conjecture seems important to me. Let $n_k < x < n_{k+1}$: then

$$J(n_k)/D(x) \to 0. \tag{14}$$

(14) would imply that (8) holds for every c. (14) seems interesting, since all our information on large values of $p_{k+1} - p_k$ comes from information on $J(n_k)$. I feel confident that (14) is true, but see no way to attack and offer 1000 dollars for any relevant information on (14).

I expect that

$$\max_{1 \le i < \phi(n_k)} \min(a_{i+1} - a_i, a_i - a_{i-1})/J(n_k) \to 0. \tag{15}$$

Perhaps (15) will not be very difficult, in any case it should be much easier than (12). (15) certainly is false for almost all integers, but may remain true for the sequence of integers satisfying $\phi(n'_k)/n'_k \to 0$ i.e., $\prod_{p/n'_k} (1 - \frac{1}{p}) \to 0$.

Is it true that if $H(n)/J(n) \to \infty$, then

$$\phi_u(x, x+H(n)) = (1+o(1)) \frac{\phi(n)}{n} H(n), \tag{16}$$

where $\phi_n(u,v)$ is the number of integers $u < m < v$ $(m,n) = 1$? (16) is related to (9) but is probably much easier. (16) certainly holds for almost all n but I can not prove it for the n_k's, but in any case I am sure if true it is much easier than (9). More than forty years ago I conjectured that if $1 \le a_1 < \ldots < a_{\phi(n)} = n - 1$ are the integers relatively prime to n, then

$$\sum_{a_i<n} (a_{1+i} - a_i)^2 < \frac{c\, n^2}{\phi(n)} \quad . \tag{17}$$

(17) was recently proved by Montgomery and Vaughan; their proof will soon appear in the Annals of Mathematics.

I thought for a while that the following conjecture is true. Let $k \geq 3$, $n_k = 2,3,\ldots p_k$. Then every even $t < J(n_k)$ is of the form $a_{\ell+1} - a_\ell$. Lacampagne and Selfridge showed that this fails for $n_6 = 300\ 30$. $J(300\ 30) = 22$ (9461 - 9439 = 22), but there is no solution for $t = 20$. In fact the conjecture may fail for all $k > k_0$. Perhaps there is an absolute constant c so that for every n the number of $t < J(n)$ which are of the form $a_{\ell+1} - a_\ell$ is $> cJ(n_k)$. It would certainly be of interest to determine or estimate the smallest even $t < J(n_k)$ not of the form $a_{\ell+1} - a_\ell$.

Let again $n_k = 2, 3,\ldots,p_k$. Let $r = r_k$ be the smallest index for which

$$a_{r+1} - a_r = J(n_k); \tag{18}$$

i.e., r is the smallest index for which $a_{i+1} - a_i$ assumes its maximum. I am sure that r increases exponentially but can not even prove that it increases faster than polynomially.

I would like to get an estimation for the number of solutions of (18), and more generally for the number of solutions of

$$a_{t+1} - a_t = s\,. \tag{19}$$

An estimation for the smallest solution of (19) would certainly be of some interest.

I conjectured some time ago that if $(a,b) = 1$, $a < b < x$ then perhaps

$$\min (J(a), J(b)) < c \log x\,. \tag{20}$$

(20) is certainly a serious conjecture, and if true or false might give some insight into the mysterious behavior of $p_{k+1} - p_k$.

Now I state some problems on sieves. An old problem of mine states as follows: Let $f(x)$ be the smallest integer so that there is a system of congruences

$$a_p \pmod p \quad p \le f(x) \tag{21}$$

so that every integer $n \le x$ satisfies at least one of the congruences (21). It is really presumptious to call this "my problem". No doubt it has been considered by many others. $f(x) > x^{\frac{1}{2}+\varepsilon}$ would have many important consequences. Perhaps $f(x) > x^{1-\varepsilon}$ holds for every $\varepsilon > 0$ if $x > x_0(\varepsilon)$.

Perhaps it is more important to study the smallest $f^{(\varepsilon)}(x)$ for which the number of integers $n < x$ which do not satisfy any of the congruences (21) is less than $\varepsilon x/\log x$. Clearly $f^{(\varepsilon)}(x) \le f(x)$. Perhaps for sufficiently small ε

$$\log f^{(\varepsilon)}(x)/\log f(x) \to 1 .$$

Let $f_r(x)$ be the smallest integer for which there is a set of r congruences

$$\begin{cases} a_{p_1} \pmod p,\ a_{p_2} \pmod p,\ \dots,\ a_{p_r} \pmod p \\ r < p < f_r(x) \end{cases} \tag{22}$$

so that every integer $n \le x$ satisfies at least one of the congruences (23). Perhaps $f_r(x) < x^{\varepsilon}$ for $r > r(\varepsilon)$, but as far as I know, $f_r(x) > x^{1-\varepsilon}$ for every r and $x > x_0(\varepsilon)$ has not been disproved.

Hildebrandt and I considered the following problem: Let $F(x)$ be the largest integer so that there is a system of congruences

$$a_p \pmod p, \quad F(x) < p < x \tag{23}$$

so that every integer $n < x$ satisfies at least one of the congruences (23). It is not hard to prove that

$$F(x) > \exp(1 - \varepsilon) \log x \log\log\log x/\log\log x$$

and very likely

$$F(x) < \exp(1+\varepsilon)\log x \log\log\log x/\log\log x .$$

A related conjecture of mine states that if we consider the congruences

$$a_p \pmod p, \quad p < x, \tag{24}$$

then for every choice of the a_p there always is an integer $n < x$ which satisfies at most one of the congruences (24). I have of course no real evidence that this is true.

Denote by $a_1^{(r)} < a_2^{(r)} < \ldots$ the set of integers which have at most r prime factors. It is a simple exercise to prove that for $r = 2$ [11]

$$\overline{\lim}\,(a_{i+1}^{(r)} - a_i^{(r)})/\log(a_i^{(r)}) > 0 . \tag{25}$$

I could never prove that the limit in (25) is ∞: also, I could get no satisfactory result for $r > 2$. The limit could very well be 0 for $r > 2$.

Now I would like to restate some old problems of Selfridge and myself [12] which seem interesting to us but which have been completely neglected, partly because our paper has been made to some extent obsolet by the results of Hensley and Richards [13]. Let

$$n < a_1 < a_2 < \ldots < a_t \le n + k, \quad (a_i, a_j) = 1, \quad 1 \le i < j \le t. \tag{26}$$

The sequence (26) is called complete if for every $n < s \le n + k$, $(s, a_i) > 1$ for some $1 \le i \le t$. Put $\max t = F(n;k)$ and $\min t = f(n;k)$ where the maximum and minimum is to be taken for all complete sequences (26). Consider the four functions

$$\max_n F(n;k), \quad \min_n F(n;k), \quad \max_n f(n;k), \quad \min_n f(n;k) .$$

Our results on $\max_n F(n;k)$ have been made obsolete by Hensley and

Richards, but perhaps it is remarkable that we could only prove

$$(27)\qquad k^{\frac{1}{2}-\varepsilon} < \min F(n;k) < c\, k(\log\log k)^2 (\log k)^{-2} (\log\log\log k)^{-1} .$$

The upper bound in (27) is clearly related to Rankin's result (8), and will be hard to improve, but the lower bound should surely be improved to $k^{1-\varepsilon}$ or at least to $k^{\frac{1}{2}+\varepsilon}$; perhaps even $\min F(n;k)/k^{\frac{1}{2}} \to \infty$ would be of some interest.

Both $\max_n F(n;k)$ and $\min_n F(n;k)$ are clearly monotonic, but $\max_n f(n;k)$ is not monotonic, since $\max_n f(n;6) = 3$ and $\max_n f(n;5) = 4$. This is the only such case we found, but we only computed $\max_n f(n;k)$ for $k \le 45$. Put

$$(28)\qquad \min_n (F(n;k) - f(n;k)) = g(k) .$$

We conjectured that $g(k) \to \infty$ as $k \to \infty$. Perhaps (28) can be proved algorithmically and will not be difficult. Clearly all the integers all whose prime factors are $\ge k$ must occur in every complete sequence. Perhaps

$$(29)\qquad \lim_{k\to\infty} \max \frac{F(n;k)}{k/\log k} > 1,$$

but as far as I know (29) is still open; we only can prove that the lim sup is finite and the lim inf ≥ 1.

It is trivial that $\min f(n;k) = 2$. Denote by n_k the smallest integer for which $f(n_k;k) = 2$. Trivially $n_k \le \prod_{p_i \le k} p_i - k$. We have a non-trivial proof that for some k there is strict inequality.

Denote further by n'_k the smallest integer for which there are two integers a and b, $n'_k < a < b < n'_k + k$ so that $(n + j, ab) > 1$ for $1 \le j \le k$. The difference between n'_k and n_k is that in the definition of n'_k we do not require $(a,b) = 1$. We show that for all sufficiently large $k < n'_k < \frac{1}{2} \prod_{p<k} p$ and probably $n'_k = o(\prod_{p\le k} p)$.

For which k is it true that if $(a,b) = 1$, $1 < b - a = k$, then there always is c, $a < c < b$ such that $(a,b,c) = 1$? Perhaps for $k > k_0$ there is no such k. If such a k exists then for this k, $n_k = \prod_{p<k} p - k$.

Is there a k so that for some set of k consecutive integers $n + 1,\ldots,n + k$

$$\left(n + i, \prod_{\substack{j=1 \\ j\neq 1}}^{k} (n + j)\right) = A(n;i)$$

is complete for every i, $1 \le i \le k$? Is there a k so that every $A(n;i)$ has more than r distinct prime factors? For $r = 0$ every sufficiently large k has this property. This is a well known result of Brauer, Pillai and Szekeres [14]. For $r > 0$ we do not know the answer, which may very well be yes for $r = 1$ and no for $r > 1$. This problem is related to (25).

In another paper, Selfridge and I [15] prove the following surprising theorem: for every $\varepsilon > 0$ and k there is a set of k^2 primes $p_1 > \ldots > p_{k^2}$ and an interval $I = \{x, x + (3 - \varepsilon)p_1\}$ so that the number of distinct integers m in I which are multiples of any of the p's is $2k$. This theorem is surprising since one would expect that the number of these integers is $> ck^2$. Since our proof is not easily accessible I give it here in full detail. First we prove that our result is best possible. In fact we show that any interval I' of length $> 2p_1$ contains at least $2k$ distinct multiples of the p's. This is essentially best possible. The interval $\{\prod_{i=1}^{k^2} p_i - p_{k^2} + 1, \prod_{i=1}^{k^2} p_i + p_{k^2} - 1\}$ has length $2p_{k^2} - 2$ and contains only one multiple of the p's. Let I' be the interval $\{a,b\}$, $b - a > 2p_1$. I'_1 is the interval $\{a, a + \frac{1}{2}(b - a)\}$ and I'_2 the interval $\{a + \frac{1}{2}(b - a), b\}$. Each of these intervals contains at least

$$\sum_{i=1}^{k^2} \left[\frac{b - a}{2p_i}\right] \ge k^2$$

multiples of the p's (counted by multiplicity). If no m in I is a multiple of more than k of the p's then clearly there are at least $2k$ distinct multiples of the p's in I. Thus, assume that there is an m in I'_1 which is a multiple of $r > k$, p's, where r is the largest such integer.

Let $p_{i_1}, \ldots, p_{i_r}$, $r > k$ be the prime factors of m. Thus in I'_1 there are at least $\frac{k^2}{r}$ distinct multiples of the p's. For every p_{i_j}, let s_j be the smallest integer for which $m + 2^{s_j} \cdot p_{i_j}$ is in I'_2.

Such an s_j clearly exists, and the numbers $m + 2^{s_j} \cdot p_{i_j}$ are clearly distinct for $j = 1,2,\ldots,r$. Thus I' contains at least $r + \frac{k^2}{r} > 2k$ distinct multiples of the p's, which completes the proof.

Now we prove the more difficult statement that there is an I of length $(3 - \epsilon)p_1$ which contains no more than 2k distinct multiples of the p's. First we prove:

LEMMA. *For every* k *and arbitrarily large* N *there are* k^2 *primes*

$$N < q_0 < q_1 < \ldots < q_{k^2-1} < N + (\log N)^{k+3}$$

satisfying for every $1 \le i \le k - 1$, $1 \le j \le k - i$

$$q_i - q_0 = q_{i+tk} - q_{tk} .$$

In other words there are k sets of k primes whose internal structure is the same. Probably very much more is true: there is an f(k) and infinitely many primes p so that all the numbers $p + t\,f(k)$, $0 \le t < k^2$, are primes --- in fact consecutive primes. Needless to say it is quite hopeless at present to prove this conjecture, and fortunately we do not need it.

The proof of the Lemma is by a simple counting argument. It follows from the prime number theorem (or a more elementary theorem) that for every large x there is an interval of length $L > (4k \log x)^{k+2}$ between $\frac{x}{2}$ and x which contains more than $\frac{L}{2 \log x}$ primes. Denote these primes by

$$y < r_1 < r_2 < \ldots < r_w < y + L, \quad w > \frac{L}{2 \log x} .$$

Consider the $[\frac{w-1}{k}]$ intervals $[r_{(u-1)k+1}, r_{uk+1}]$, $uk + 1 < w$. We only retain those intervals which are shorter than 4k log x. Clearly there are at least $L(4k \log x)^{-1}$ such intervals. The number of patterns for the k primes $r_{(u-1)k+1}, r_{(u-1)k+2}, \ldots, r_{uk}$ in these intervals is clearly less than $(4k \log x)^{k+1}$. Thus, for sufficiently large x, there are more than k k-tuples of primes giving the same pattern, which completes the proof of our Lemma.

Now using the Chinese remainder theorem we are ready to complete the proof of our theorem. Put

$$\alpha_i = \prod_{j=0}^{k-1} q_{ik+j}, \quad \beta_j = \prod_{i=0}^{k-1} q_{ik+j}, \quad 1 \le i, \ j \le k-1.$$

Clearly

$$\prod_{i=0}^{k-1} \alpha_i = \prod_{j=0}^{k-1} \beta_j = \prod_{\ell=0}^{k^2-1} q_\ell .$$

Now we determine $x \bmod \prod_{\ell=0}^{k^2-1} q_\ell$ as follows:

$$x + q_j \equiv 0 \ (\mathrm{mod}\ \beta_j), \quad x + q_0 \equiv q_{jk} (\mathrm{mod}\ \alpha_j), \quad 0 \le j \le k-1.$$

A simple argument shows that the interval $\{x - q_0 + 1, x + 2q_0 - 1\}$ of length $3q_0 - 2 > (3 - \varepsilon) q_{k^2-1}$ contains only $2k$ multiples of the q's: namely, the unique multiples of $\alpha_0, \alpha_1, \ldots, \alpha_{k-1}; \beta_0, \beta_1, \ldots, \beta_{k-1}$.

Let now again $p_1 > p_2 > \ldots > p_{k^2}$, and I is an interval of length $\ge 3p_1$. Unfortunately, here, so to speak, "all hell breaks loose," and we completely loose control over the distinct multiples of the p's. It is quite possible that in this case, I contains more than $c\,k^2$ distinct multiples of the p's. I can only prove the following much weaker theorem.

Let $p_1 > \ldots > p_{k^2}$, and I be an interval of length $\ge 3p_1$. Then I contains at least $6^{\frac{1}{2}}k$ distinct multiples of the p's.

Clearly the interval I contains at least $3\,k^2$ multiples of the p' counted by multiplicity. Let r be the largest integer so that there is an m in I which is the multiple of r p's; say, $m \equiv 0 \ (\mathrm{mod}\ p_{\ell_1}, \ldots, p_{\ell_r})$.

Each p_{ℓ_j}, $j = 1, \ldots, r$ has at least two other multiples in I (namely $m \pm p_{\ell_j}$ or $m + p_{\ell_j}$, $m + 2p_{\ell_j}$ or $m - p_{\ell_j}$, $m - 2p_{\ell_j}$). These $2r + 1$ multiples of the p's are clearly all distinct. Thus I contains at least

$$\min\left(\frac{3k^2}{r}, \ 2r + 1\right) > 6^{\frac{1}{2}}k$$

distinct multiples of the p's, which completes our proof of our theorem. I am sure that this result is not the best possible. Perhaps the following related problem is also interesting: determine the smallest f(u)

so that if $p_1 > \ldots > p_u$ are primes, every interval of length $f(u)p_1$ contains an integer divisible by precisely one of the p's. Clearly many related questions can be asked.

Denote by I_n the interval $(\frac{n}{3},\frac{n}{2})$ and by $f(x,n)$ the number of integers m, $x < m < x + n$ which have at least one prime factor in I_n. An old conjecture of mine states

$$f(x,n) > cn/\log n. \tag{30}$$

It seems ridiculous that I have not been able to make any progress with (30) and I am not sure if I am just being silly and missing an obvious point, or whether (30) is really difficult or at least requires a clever idea. It is easy to see that the number of integers having at least two prime factors in $\{x,x + n\}$ is at most

$$\tfrac{1}{2}(\pi(\tfrac{n}{2}) - \pi(\tfrac{n}{3})) = (1 + o(1))\frac{n}{12 \log n},$$

and that equality is possible here: also $f(x,n) \le 2(\pi(\frac{n}{2}) - \pi(\frac{n}{3}))$ for suitable values of x, and equality is again possible, but I could only prove $f(x,n) > c(\frac{n}{\log n})^{\frac{1}{2}}$. It is not difficult to show that there is an absolute constant C so that if $n \to \infty$ then for almost all x

$$f(x,n) = (C + o(1))\frac{n}{\log n},$$

and with a little more trouble one could obtain results on the distribution function of the error $f(x,n) - C\frac{n}{\log n}$. None of this seems to help with (30).

To finish the paper let me just state a few older problems. Denote by $p_1,p_2,\ldots$ the sequence of primes. Prachar and I [16] conjectured that the number of indices k for which for every $i < k < j$,

$$p_i/i < p_k/k < p_j/j, \tag{31}$$

is finite.

(31) seems very plausible and it probably holds for many other sequences; e.g., for the primes $p \equiv a \pmod b$ or for the set of integers

not divisible by a sec of primes $\Sigma\ 1/p_i = \infty$, where the complementary set q_i also satisfies $\Sigma\ 1/q_i = \infty$. In fact (31) should hold if $a_k/k \to \infty$, but not too fast, and a_k is not too regular. These rather vague statements, of course, do not really help, and it must be left open whether any non-trivial statement related to (31) can be made and proved.

More than twenty-five years ago I made the following (foolish) conjecture. Let $a_1 < a_2 < \ldots < a_k \le n$, $\prod_{i=1}^{k} 1/a \le 1$. Is it then true that the number of integers not exceeding n which are not divisible by any of the a's is $> cn$? This was disproved by Schinzel and Szekeres [17] and more recently Ruzsa and Tenenbaum proved that the number of these integers is $> c_1 \frac{n}{\log n}$, but can be less than $c_2 n/\log n$.

Let $p_1 < p_2 < \ldots < n$ be a sequence of primes for which $\Sigma\ 1/p_i \le 1$. Then it is easy to see that there are cn integers, no one of which is a multiple of any of the p's $\le n$. It will perhaps not be difficult to determine the smallest possible value of c.

One of the most interesting unconventional problems of primes is due to Ostman: prove that one can not find two sequences $a_1 < a_2 < \ldots$, $b_1 < b_2 \ldots$ of at least two elements, so that all but a finite number of primes are of the form $a_i + b_j$ and only a finite number of composite numbers are of the form $a_i + b_j$, in other words, the symmetric difference of the primes and the integers of the form $a_i + b_j$ must be infinite. This striking conjecture is still open. Hornfeck [18] proved it in the case that one of the sequences $a_1 < a_2 < \ldots$ or $b_1 < b_2 < \ldots$ is finite.

It follows from the prime k-tuple conjecture that there are two infinite sequences $a_1 < a_2 < \ldots$, $b_1 < b_2 < \ldots$ so that all the sums $a_i + b_j$ are primes. It seems certain that at least one of these sequences must tend to infinity at least exponentially. By the way, it seems certain that if there are only a finite number of composite numbers among the $a_i + b_j$ then there are only $o(\frac{x}{\log x})$ primes $p < x$ of the form $a_i + b_j$, which would be much stronger than Ostman's conjecture. Since the analog of the prime k-tuple conjecture clearly holds for the squarefree numbers, it is easy to see that there are infinite sequences $a_1 < a_2 < \ldots$, $b_1 < b_2 < \ldots$ so that all the integers $a_i + b_j$ are squarefree. Perhaps it is true that if all but a finite number of the $a_i + b_j$ are squarefree, and both sequences a_i and b_j are infinite, then the number of squarefree integers of the form $a_i + b_j$ is $o(x)$, or even

lightly stronger; $A(x)B(x) = o(x)$ where $A(x) = \sum_{a_i<x} 1$, $B(x) = \sum_{b_i<x} 1$. Pomerance once asked: is there a subsequence of the primes $p_{i_1} < p_{i_2} < \dots$ whose second difference $p_{i_r} - 2p_{i_{r+1}} + p_{i_{r+2}}$ is bounded from above (or bounded in absolute value). Probably such a sequence does not exist, not even if the primes are replaced by the squarefree numbers, but I do not see how to attack these questions. About thirty years ago, Ricci and I [19] proved that the set of limit points of $(p_{k+1} - p_k)/\log k$ is of positive Lebesgue measure. Unfortunately ∞ is the only limit point of this set known to us. Can one prove that this set has a finite limit point ≥ 1?

Perhaps the following somewhat vague conjecture is not hopeless: let $H(x)/\log x \to \infty$ smoothly but $H(x) < L(x)$ (see (8)). Is it then true that the set of limit points of $(p_{k+1} - p_k)/H(k)$ has positive measure? Is there for every C an index k for which

$$C \log x < p_k - p_{k-1} < p_{k+1} - p_k, \quad p_k < x?$$

Finally I state a somewhat unconventional problem which was considered by Pomerance and myself. Straus and I once conjectured that if $k > k_0$, then there always is an i for which

$$p^2_k < p_{k+i}p_{k-i} . \tag{32}$$

Pomerance [20] disproved this: in fact he disproved this for much more general sequences. We tried unsuccessfully to prove that in fact for almost all k (32) holds. It would suffice to show, that for almost all k there is an i for which

$$2p_k > p_{k+i} + p_{k-i}, \quad p_{k+i} < p_k + p_k^{\frac{1}{2}} , \tag{33}$$

but we could not prove (33). Is it true that the number of distinct integers of the form $p_{n+i} + p_{n-i}$, $i = 1,2,\dots$ is $> cn/\log n^2$? It easily follows from the sharper form of the prime number theorem, that the number of solutions of $A = p_{n+i} + p_{n-i}$ in i is bounded if $n \to \infty$, but we can show this only for the A's in the neighborhood of $2p_n$.

Pomerance and I further considered the following problems: is it true that for $n > n_0$ there always is an i for which $2p_n = p_{n+i} + p_{n-i}$? The answer is almost certainly affirmative. Is it true that there is a c so that for infinitely many i and every $i < n$

$$p_{n+i} + p_{n-i} - 2p_n > -C?$$

Put

$$M(n) = \max_i p_{n+i}p_{n-i} .$$

Is it true that there is an $\alpha > 0$ so that for infinitely many n

$$M_n > p_{n+i}p_{n-i} + n^{\alpha}? \tag{34}$$

If the answer is affirmative, try to determine the largest α for which (34) holds for infinitely many n.

Finally, I would like to remark that (17) leads to interesting and deep problems for other sequences; e.g., let $q_1 < q_2 < \dots$ be the sequence of consecutive squarefree numbers. Is it true that for every α

$$\sum_{q_n<x} (q_{n+1} - q_n)^{\alpha} < c_{\alpha}x? \tag{35}$$

I proved (35) for every $\alpha \le 2$ and Hooley [21] proved it for every $\alpha \le 3$. (Hooley just informed me that he can prove it for every $\alpha \le 3 + \varepsilon$ for some small positive ε.) If (35) holds for every α, then for every $\varepsilon > 0$ and $n > n_0(\varepsilon)$, $q_{n+1} - q_n < q_n^{\varepsilon}$. Thus (35), if true, is probably very deep. I could not disprove the following much stronger conjecture

$$\sum_{q_n<x} \exp C(q_{n+1} - q_n) < \alpha_C x. \tag{36}$$

(34), if true, is certainly beyond our reach, but perhaps (36) can be disproved.

REFERENCES

1. P. Erdös and L. Mirsky, *On the distribution of values of the divisor function* d(n), Proc. London Math. Soc. 3 (1952), 257-271.

2. P. Erdös, *Some remarks on Euler's ϕ-function*, Acta Arith. 4 (1958), 10-19.

3. D. R. Heath-Brown, *The divisor function at consecutive integers*, Mathematica 31 (1984), 141-149.

4. For a more detailed proof of these statements see a forthcoming paper of Pomerance, Sárközy and myself.

5. H. Cramer, *On the order of magnitude of the difference between consecutive prime numbers*, Acta Arith. 2 (1937), 23-46.

6. R. A. Rankin, *The difference between consecutive prime numbers*, J. London Math. Soc. 13 (1938), 242-247.

7. See P.D.T. A. Elliott, Probabilistic Number Theory, Springer-Verlag 1980, Vol. 1, p. 254.

8. H. Maier, Michigan Math. Journal 32 (1985), 221-225.

9. P. Erdös, *On the integers relatively prime to* n *and on a number theoretic function considered by Jacobstahl*, Math. Scand. 10 (1962), 163-170. For many further problems and results see, C. Hooley, *On the difference of consecutive numbers relatively prime to* n, *I, II and III*, Acta Arithmetica 8 (1963), 343-347; Publ. Math. Delrecen 12 (1965), 39-49; Math. Zeitschrift 90 (1965), 355-369.

10. Iwaniec, *On the problem of Jacobstahl*, Demonstratio Mathematica 11 (1978), 1-7.

11. P. Erdös, *Problem 237*, Elemente der Mathematik 10 (1955).

12. P. Erdös and T. Selfridge, *Complete prime subsets of consecutive integers*, Proc. Manitoba Conf. 1971, 1-14.

13. Hensley and I. Richards, *Primes in intervals*, Acta Arithmetica 25 (1973), 375-391, see also P. Erdös and I. Richards, *Density functions for prime and relatively prime numbers*, Monatshefte Math. 83 (1977), 99-112.

14. R. I. Evans, *On blocks of consecutive integers*, Amer. Math. Monthly 76 (1969), 48-49.

15. P. Erdös, *Problems and results in combinatorial analysis and combinatorial number theory,* Proceedings of the Ninth Southeastern Conference on Combinatorics Graph Theory and Computing 1978, 29-40. See also a related problem of Suranyi and myself: P. Erdös and J. Suranyi, *Remarks on a problem of a mathematical competition* (in Hungarian) Math. Lapok X (1959), 39-47.

16. P. Erdös and K. Prachar, *Sätze un Probleme über* $p_{h/k}$. Abhandlungen Math. Sem. Hamburg 25 (1961-62), 251-256.

17. A. Schinzel and Szekeres, *Sur un probleme de Paul Erdös,* Acta Sci. Math. Szeged 20 (1959), 221-229.

18. B. Hornbeck, *Ein Satz über die Primzahlmenge,* Math. Zeitschrift 60 (1954), 271-273.

19. G. Ricci, *Recherches sur l'allure de la suite* $p_{n+1} - p_n / \log p_n$, Colloque sur la théorie des nombres, Bruxelles (1955), 93-106. Liège and Paris. My paper appeared in the lecture notes of a conference held in 1955 at Lake Como, Italy. This paper has not been reviewed and is not easily accessible.

20. C. Pomerance, *The prime number graph,* Math. Comp. 33 (1979), 399-408.

21. C. Hooley, *On the distribution of squarefree numbers,* Canad. J. Math. 25 (1973), 1216-1223.