Some Problems in Number Theory

P. Erdős

Academy of Sciences, Budapest, Hungary

In the present paper I discuss some problems in number theory which I have thought about in the last few years; computational techniques can be applied to some of them.

1. On Prime Factors of Consecutive Integers

Let $f(k)$ be the smallest integer with the property that the product of $f(k)$ consecutive integers all greater than k is always divisible by a prime greater than k. A well-known theorem of Sylvester and Schur (see Erdős, 1934) states that $f(k) \leqslant k$. I proved (1955)

$$c_1 \log k \log_2 k \log_4 k/(\log_3 k)^2 < f(k) < c_2 k/\log k,\dagger$$

Recently Ramachandra (1969 and to appear) proved $f(k) < (1 + o(1))\, k/\log k$. It seems to me to be very difficult to prove that for all $k > k_0$ we have $f(k) < \pi(k)$, though I have no doubt that the conjecture is true. In fact it seems likely that $f(k)$ is not substantially larger than

$$A_k = \max(p_{r+1} - p_r), \qquad k < p_r < p_{r+1} < 2k.$$

In fact I cannot even disprove $f(k) = A_k$ for all sufficiently large k, though it seems likely that $f(k) > A_k$ for all large k. A well known theorem of Pólya and Störmer states that if $u > u_0(k)$ then $u(u+1)$ always contains a prime factor greater than k, thus $f(k)$ can be determined in a finite number of steps, and an explicit bound has been given by Lehmer (1964) for the number of necessary steps. It is known (Utz, 1961) that $f(2) = 2, f(3) = f(4) = 3$, $f(5) = \ldots = f(10) = 4$.

Selfridge and I conjected that if $m \geqslant 2k$ then $\binom{m}{k}$ has a prime factor $\leqslant m/2$, the only exception being $\binom{7}{3}$. This conjecture was recently proved by Earl Ecklund.

† We write $\log \log k = \log_2 k$, etc.

Selfridge and I proved that there is an absolute constant $c > 0$ so that if $m \geqslant 2k$ then $\binom{m}{k}$ always has a prime factor less than m/k^c.

The proof is very simple. Assume first $m \geqslant 2k^{1+c}$, put $l = [k^c] + 1$. It follows from the theorem of Hoheisel–Ingham (see Ingham, 1937) that for sufficiently small $c > 0$ there is a prime p satisfying

$$\frac{m}{l} > p > \frac{m-k}{l} > k.$$

Clearly this prime divides $\binom{m}{k}$ and this proves our assertion if $m \geqslant 2k^{1+c}$. Assume next $m < 2k^{1+c}$. Let

$$s = \left[\frac{3m}{2k}\right] + 1.$$

It follows from the Hoheisel–Ingham theorem that there is a prime p satisfying

$$\frac{m}{s} > p > \frac{m-k}{s-1}.$$

Clearly

$$p \Big| \binom{m}{k} \qquad \left(\text{since } \frac{k}{2} < \frac{m-k}{s-1} < p < k \text{ amd } m - k < (s-1)p < sp < m\right)$$

which completes our proof. The simplicity of our proof is caused by the fact that we have not determined c explicitly.

Selfridge and I conjectured that if $m > k^2$ then $\binom{m}{k}$ has a prime factor $\leqslant m/k$; $\binom{7}{3}$ is certainly an exception, and this may be the only one. In connection with this problem we asked: Determine or estimate the smallest integer $g(k)$ so that all prime factors of $\binom{g(k)}{k}$ are greater than k, (it is easy to see that such integers exist).

It is perhaps true that, for $k > k_0(\varepsilon)$ and $m > k^{1+\varepsilon}$, $\binom{m}{k}$ always has a prime factor greater than $k^{1+\varepsilon} - k$. Ramachandra (1969) has some results which point in this direction. More generally let $h(k)$ be the largest integer so that if $m > h(k)$ then $\binom{m}{k}$ always has a prime factor greater than $h(k) - k$.

I am sure that $h(k) > k^c$ for every $c > 0$ and $k > k_0(c)$; $h(k) > ck \log k$ is easy and Ramachandra's result will no doubt give $h(k) > (1 + o(1))\, k \log k$. Denote by p_k the least prime greater than $2k$. Faulkner (1966) proved that for $m \geqslant p_k$, $\binom{m}{k}$ always has a prime factor $\geqslant p_k$, except for $\binom{9}{2}$ and $\binom{10}{3}$. Thus $h(k) \geqslant p_k + k$ for $k > 3$.

It is easy to see that $h(2) = 4$, $h(3) = 6$, $h(4) = 16$ (i.e. the product of 4 consecutive integers $\geqslant 13$ always has a prime factor $\geqslant 13$). It is difficult to compute $h(k)$ but by the effectivisation results of Brown this can be done in a bounded number of steps. Lehmer (1963) showed $h(7) \geqslant 43$.

I conjectured that, for every $m \geqslant 2k$, $\binom{m}{k}$ has a divisor d with $m - k < d \leqslant m$. This is easy to see if $k = p^z$. Schinzel (1958) proved that in general it is incorrect, e.g., it is false for $k = 15$, $m = 99125$. He further proved that it is true for all integers $k \leqslant 34$ except 15, 21, 22, 33. Schinzel now conjectures that it is false for all $k > 34$, $k \neq p^z$. This conjecture has been verified for $k < 150$. I proved (see Schinzel, 1958) that my conjecture is false for infinitely many $k \neq p^z$.

In view of the failure of my conjecture one can try to investigate the greatest factor of $\binom{m}{k}$ not greater than m. I would now conjecture that the greatest prime factor $\leqslant m$ of $\binom{m}{k}$ is greater than cm for some $c > 0$. Unfortunately I can prove no non-trivial result.

This question leads me to the following one: Is it true that for every $\varepsilon > 0$ there is a k_0 so that, for $k > k_0$, $k!$ is the product of k integers all greater than $(k/e)(1 - \varepsilon)$. It easily follows from Stirling's formula that if

$$k! = \prod_{i=1}^{k} a_i, \quad a_1 \leqslant \ldots \leqslant a_k,$$

then $a_1 < k/e$, thus our conjecture if true is best possible.

Recently Selfridge and I proved that the product of consecutive integers is never a power (that it is never a square is due to Rigge, 1939); our proof is not quite easy and will be published elsewhere (for a weaker result see Erdős, 1955b). In fact we prove a somewhat stronger result. We prove that for every $l > 1$, $k > 1$ the product $\prod_{i=1}^{k} (m + i)$ contains a prime $p > k$ to an exponent which is not a multiple of l. We conjecture that if $l \geqslant 2$ and $k \geqslant 3$ then $\prod_{i=1}^{k} (m + i)$ contains a prime $p > k$ to the exponent one. The only exception is $48 \cdot 49 \cdot 50$. For $k = 2$ there are infinitely many exceptions. This conjecture if true is very deep.

Put ($p^z \| m$ means $p^z | m$, $p^{z+1} \nmid m$)

$$A_i^{(m)} = \Pi p^z, \quad p^z \| (m+i), \quad p \leqslant k.$$

It is not difficult to prove that for $k > k_0(\varepsilon)$

$$\min_{1 \leqslant i \leqslant k} A_i^{(m)} < (1+\varepsilon)k.$$

Probably very much more is true, in fact perhaps

$$\lim_{k \to \infty} \frac{1}{k} \max_{0 \leqslant m < \infty} \min_{1 \leqslant i \leqslant k} A_i^{(m)} = 0.$$

2. Covering Congruences

A system of congruences $a_i \pmod{m_i}$, $m_1 < \ldots < m_k$ is called a covering system if every integer satisfies at least one of the congruences $a_i \pmod{m_i}$. I was lead to the problem of covering congruences by a letter of Romanoff who asked if there are infinitely many odd integers not of the form $2^k + p$ (as is well known Romanoff (1934) proved that the lower density of the integers of the form $2^k + p$ is positive).

The simplest covering system is 0 (mod 2), 0 (mod 3), 1 (mod 4), 1 (mod 6), 11 (mod 12) and the system 0 (mod 2), 0 (mod 3), 1 (mod 4), 7 (mod 8), 11 (mod 12), 19 (mod 24) shows (Erdős 1947-51) that the answer to Romanoff's question is positive, in fact there is an arithmetic progression consisting entirely of odd numbers no term of which is of the form $2^k + p$.

The following question seems very difficult: Is it true that to every c there exists a covering system $a_i \pmod{m_i}$ $c \leqslant m_1 < \ldots < m_k$? This is known for $c \leqslant 9$ (see Churchhouse, 1968) but the general case seems very difficult. A positive answer would imply that for every r there is an arithmetic progression no term of which is the sum of a power of 2 and an integer having at most r prime factors.

Schinzel recently investigated the question whether, for fixed r, there is an arithmetic progression no term of which is of the form $2^{k_1} + 2^{k_2} + \ldots + 2^{k_r} + p$; already for $r = 2$ the question seems difficult. Schinzel (1967) recently applied covering congruences to the study of reducibility of polynomials.

There are many further interesting problems on covering congruences, e.g., is there a covering congruence all whose moduli are odd, or is there a covering congruence in which no two moduli divide each other? Schinzel (1967) and Selfridge observed that the two problems are connected.

Call an integer m covering if one can find a covering set whose moduli are all divisors of m; $m = 12$ is clearly the smallest covering integer. Clearly all multiples of a covering integer are again covering. An integer is primitive

covering if it is covering but all its divisors are not covering. Clearly we obtain the covering integers by taking the set of all multiples of the primitive covering integers. I can prove using the results in Erdős (1948) that the covering integers have a density. One could try to estimate the number of primitive covering integers not exceeding x.

I expect that for every $c > 0$ there is an m which is not covering and for which $\sigma(m)/m > c$, but I could not prove this (perhaps I overlook a simple idea).

A system of arithmetic progressions $a_i \pmod{m_i}$, $m_1 < \ldots < m_k$ is called disjoint if no integer is in two of them. Denote by $f(x)$ the maximum number of pairwise disjoint arithmetic progressions whose difference does not exceed x. Stein and I conjectured that $f(x) = o(x)$; Szemeredi and I (1968) recently proved this. The sharpest results for $f(x)$ are

$$x \exp(-c_1(\log x \log_2 x)^{\frac{1}{2}}) < f(x) < x(\log x)^{-c_2},$$

perhaps the lower bound is close to the true order of magnitude.

Stein conjectured that if $a_i \pmod{m_i}$, $m_1 < \ldots < m_k$ are k disjoint congruences there is an integer $\leqslant 2^k$ which does not satisfy any of these congruences Selfridge proved this conjecture. I conjectured that if $a_i \pmod{m_i}$, $m_1 < \ldots < m_k$ are any k congruences which are not covering then there is an integer $\leqslant 2^k$ which does not satisfy any of these congruences (Selfridge, Crittenden and Van der Eyden recently proved this conjecture).

It is not hard to see that the density of integers not satisfying any of the disjoint congruences $a_i \pmod{m_i}$, $m_1 < \ldots < m_k$ is $\geqslant 1/2^k$ and that this result is best possible. The same result probably holds for any k congruences which are not covering (Erdős, 1962).

I would like to state one more problem on arithmetic progressions: Let $a_i \pmod{m_i}$, $m_1 \leqslant m_2 \leqslant \ldots$ be an infinite sequence of arithmetic progressions. Is it true that the set of integers not satisfying any of these congruences always has a logarithmic density? Special cases of this conjecture were proved by Davenport and myself (1936 and 1951).

3. Some Problems and Results on the Addition of Residue Classes

Heilbronn and I (1969) proved that if $a_1, \ldots, a_k$, $k \geqslant 3(6p)^{\frac{1}{2}}$ are distinct residues mod p (p prime) then every residue (mod p) can be written in the form

$$\sum_{i=1}^{k} \varepsilon_i a_i, \quad \varepsilon_i = 0 \text{ or } 1.$$

We conjectured that the same holds for $k > 2\sqrt{p}$ and that this is best possible. Olsen (1968) recently proved this conjecture. We further conjectured

that the number of distinct residues of the form $a_i + a_j$, $1 \leqslant i < j \leqslant k$, is at least $2k - 3$; as far as I know this conjecture is still unsettled.

Let now m be composite and $a_1, \ldots, a_k$ be k distinct residues mod m. We conjectured (Erdős and Heilbronn, 1969) that if $k > c\sqrt{m}$ then

$$\sum_{i=1}^{k} \varepsilon_i a_i \equiv 0 \pmod{m}, \qquad \varepsilon_i = 0 \text{ or } 1$$

is always solvable (probably $k > \sqrt{2m} + o(\sqrt{m})$ will suffice). Ryavec (1968) proved a slightly weaker result and our conjecture was recently proved by Szemeredi (his paper will appear in *Acta Arithmetica*). Szemeredi's proof works for every abelian group of order m; perhaps the result holds for non-abelian groups too.

Eggleston proved the following result: Let G_m be an abelian group of m elements, $m \leqslant n + k - 1$, $a_1, \ldots, a_n$ are n elements of G_m where at least k of the a's are distinct. Then (e is the unit element of G_m)

$$e = \prod_{i=1}^{k} a_i^{\varepsilon_i}, \quad \varepsilon_i = 0 \text{ or } 1$$

is always solvable.

Eggleston and I conjectured that $m \leqslant n + k - 1$ can be replaced by $m \leqslant n + \binom{k}{2}$; this if true is easily seen to be best possible (it suffices to take G_m to be the additive group mod m and the a's $1, \ldots, k, 1, \ldots, 1$).

We proved this conjecture if $m > m_0(k)$ (unpublished), also we were led to the following question which seems to be of some interset. Let $f(k)$ be the largest integer with the following property; let $a_1, \ldots, a_k$ be k distinct elements of G_m and assume that no product

$$\prod_{i=1}^{k} a_i^{\varepsilon_i}, \quad \varepsilon_i = 0 \text{ or } 1,$$

equals the unit of G_m; then at least $f(k)$ distinct elements of G_m can be represented in the form

$$\prod_{i=1}^{k} a_i^{\varepsilon_i}, \quad \varepsilon_i = 0 \text{ or } 1.$$

We showed $f(2) = 2, f(3) = 5, f(4) = 8, f(k+1) \geqslant f(k) + 2$. Szemeredi showed $f(k) > ck^2$. It does not seem to be easy to determine $f(k)$ or even to give an asymptotic formula for it. These problems can be stated for non-abelian groups too.

4. Miscellaneous Problems, Results and Conjectures

Denote by $\pi(x)$ the number of primes not exceeding x. Is it true that $\pi(x+y) \leqslant \pi(x) + \pi(y)$)? This conjecture, if true, is certainly extremely deep. It is not hard to prove for small values of y. I do not know for how large values of y it has been proved and I also do not know for how large values it has been checked.

Following Hardy and Littlewood (1923) put

$$\rho(y) = \lim_{x=\infty} \sup \left(\pi(x+y) - \pi(x)\right)$$

Probably $\lim_{y=\infty} \rho(y) = \infty$. Hardy and Littlewood conjectured that for $y > y_0$ then $\rho(y) > y/\log y$; this if true is certainly very deep. Using Brun's method they proved $\rho(y) < cy/\log y$ (as far as I know this is the only time they used Brun's method). Denote by $h_m(k)$ the number of integers $m < x \leqslant m + k$ which are not divisible by any prime less than or equal to k. Hardy and Littlewood conjectured that $\rho(k) = \max_m h_m(k)$. It seems probable that $\lim_{y=\infty} \left(\pi(y) - \rho(y)\right) = \infty$.

All these conjectures seem hopeless at present. Perhaps the following questions deserve some investigation. A sequence $m < a_1 < \ldots < a_l \leqslant m + k$ is called complete if $(a_i, a_j) = 1$, $1 \leqslant i < j \leqslant l$, but for every $m < n \leqslant m + k$, $(n, a_j) > 1$ for some $1 \leqslant j \leqslant l$. Denote by $f(m, k)$, respectively, $F(m, k)$ the smallest (largest) value of l. It is easy to see that $\min_m f(m, k) = 2$ $(m = k! - 1)$ but it seems very difficult to determine or give a good estimation for $\max_m f(m, k)$, $\min_m F(m, k)$ or $\max_m F(m, k)$. Clearly all three functions tend to infinity with k, perhaps $\max_m F(m, k) = \pi(k) + 1$ (clearly $\max_m F(m, k) \geqslant \pi(k) + 1$, to see this observe that the $\pi(k) + 1$ integers $k! + 1$, $k! + p$ [p runs through the primes not exceeding k] are pairwise relatively prime). $F(m, k) < ck/\log k$ trivially follows from Brun's method. For small values of k it is easy to compute all these functions.

One could try to estimate $f(m, k)$ and $F(m, k)$ if both m and k tend to infinity e.g. is it true that if c is a sufficiently large constant then $f(m, (\log m)^c)$ tends to infinity together with m? This question is connected with the problem of the difference of consecutive primes and seems very difficult.

The sharpest known inequality for large differences of consecutive primes is due to Rankin (1938) and states that for infinitely many n we have

$$p_{n+1} - p_n > c \log p_n \log_2 p_n \log_4 p_n / (\log_3 p_n)^2.$$

Denote now by $a_1^{(r)} < a_2^{(r)} < \ldots$ the sequence of integers which have at

most r prime factors. I proved (Erdős, 1955c, 1956)

$$\limsup_{k=\infty} (a_{k+1}^{(2)} - a_k^{(2)})/\log k > c;$$

perhaps this inequality holds for every r and perhaps the lim sup is in fact infinite, but I cannot prove this even for $r = 2$.

Let $g(m)$ be the smallest integer so that at least one of the integers m, $m+1, \ldots, m+g(m)$ divides the product of the others. It is easy to see that $g(k!) = k$ and, for $m > k!$, $g(m) > k$. I can prove that for infinitely many m

$$g(m) > \exp\left((\log m)^{\frac{1}{2}-\varepsilon}\right).$$

I have no good upper bound for $g(m)$. $g(m) < c\sqrt{m}$ is easy but probably $g(m) = O(m^{\varepsilon})$ and in fact perhaps $g(m) = O\left(\exp\left((\log m)^{\frac{1}{2}+\varepsilon}\right)\right)$.

Denote by $u_1^{(\varepsilon)} < \ldots < u_s^{(\varepsilon)} \leqslant m$ the integers not exceeding m all whose prime factors are $< m^{\varepsilon}$. $g(m) = O(m^{\varepsilon})$ would follow if we could show $a_{i+1} - a_i = O(m^{\varepsilon})$, but this seems hopeless at present.

Put $f(m) = \sum_{p \mid m} p$ (this function has recently been investigated from a different point of view by Mohan Lal, 1969). Denote by $F(x)$ the number of distinct integers of the sequence $f(m)$, $1 \leqslant m \leqslant x$. I can prove (unpublished)

$$c_1\, x/\log x \prod_{k=3}^{r} \log_k x < F(x) < c_2\, x/\log x \prod_{k=3}^{r} \log_k x, \tag{4.1}$$

where $1 \leqslant \log_r x \leqslant e$. Analogous questions have been investigated for the functions $\sigma(m)$, $\phi(m)$ and $d(m)$; see Erdős (1935, 1945) and Erdős and Mirsky (1952).

The same function which appears in (4.1) occurs in a completely different question. Let $1 \leqslant a_1 < \ldots < a_k \leqslant x$ be a sequence of integers so that all the sums

$$\sum_{i=1}^{k} \varepsilon_i/a_i, \quad \varepsilon_i = 0 \text{ or } 1,$$

are all different. Put $\max k = f(x)$. Then

$$c_1\, x/\log x \prod_{k=3}^{r} \log_k x < f(x) < c_2\, x \log x \prod_{k=3}^{r} \log_k x. \tag{4.2}$$

The proof of (4.2) is not published.

Finally I state a conjecture of the 16-year old Hungarian mathematician I. Ruzsa.

Let $f(m)$ be a multiplicative function whose values are elements of a group G. Let g be an element of this group, Is it true that the density of integers m for which $f(m) = g$ always exists? This conjecture if true must be very deep since it would imply the theorem of Wirsing (1967) that every multiplicative function which only assumes the values < 1 has a mean value.

References

Churchhouse, R. F. (1968). "Covering sets and systems of Congruences in Computers in Mathematical Research". pp. 20–36, North Holland.

Davenport, H. and Erdős, P. (1936). On sequences of positive integers, *Acta Arith.* **2,** 147–151 and (1951) *J. Ind. Math. Soc.* **15,** 19–24.

Erdős, P. (1934). On a theorem of Sylvester and Schur. *J. London Math. Soc.* **9,** 282–288.

Erdős P. (1935). On the normal number of prime factors of $p - 1$ and some related problems concerning Euler's ϕ function, *Quarterly J. Math.* **6,** 205–213.

Erdős, P. (1945). Some remarks on Euler's ϕ function and some related problems. *Bull. Amer. Math. Soc.* **51,** 540–544.

Erdős, P. (1947–51). On integers of the form $2^k + p$ and some related problems. *Summa Brasil. Math.* **2,** 113–123.

Erdős, P. (1948). On the density of some sequences of integers. *Bull. Amer. Math. Soc.* **54,** 685–692.

Erdős, P. (1955a). On consecutive integers. *Nieuw Archiev Wiskunde,* **3,** 124–128.

Erdős, P. (1955b). On the product of consecutive integers, III. *Indag. Math.* **17,** 85–90.

Erdős, P. (1955c). *Elemente der Mathematik,* Vol.I., 47.

Erdős, P. (1956). *Elemente der Mathematik,* Vol. II., 86–88.

Erdős, P. (1962). Szamelmeleti megjegyzesek IV. (in Hungarian), *Mat. Lapok* **13,** 241–242.

Erdős, P. and Heilbronn, H. (1969). On the addition of residue classes mod p. *Acta Arith.* **9,** 149–159.

Erdős, P. and Mirsky, L. (1952). The distribution of values of the divisor function $d(m)$. *Proc. London Math. Soc.* **3,** 257–271.

Erdős, P. and Szemeredi, E. (1968). On a problem of Erdős and Stein. *Acta Arith.* **15,** 85–90.

Faulkner, M. (1966). On a theorem of Sylvester and Schur. *J. London Math. Soc.* **41,** 107–110.

Hardy, G. H. and Littlewood, J. E. (1923). Some problems on partitio numerorum III: On the expression of a number as a sum of primes. *Acta Math.* **44,** 1–70.

Ingham, A. E. (1937). On the difference between consecutive primes. *Quarterly J. Math.* **8,** 255–266.

Lal, M. (1969). Iterates of a number theoretic function. *Math of Computation.* **23,** 181–183.

Lehmer, D. H. (1963). Some high speed logic. *Proc. Symp. in Applied Math.* **15,** 141–145.

Lehmer, D. H. (1964). On a problem of Störmer. *Illinois J. Math.* **8,** 57–79.

Olsen, J. E. (1968). An addition theorem modulo p. *J. Combinatorial Theory* **5,** 45–52.

Ramachandra, K. (1969). A note on numbers with a large prime factor. *J. London Math. Soc.* (2), **1,** 303–306.

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

Rigge, S. (1939). Über ein diophantisches Problem, 155–160. 9th Scandinavian Maths. Congress.

Romanoff, N. P. (1934). Über einige Sätze der additiven Zahlentheorie, *Math. Annalen,* **109,** 668–678.

Ryavec, C. (1968). On the addition of residue classes modulo *n*, *Pacific J. Math.*, **26,** 367–373.

Schinzel, A. (1958). Sur un probleme de P. Erdős, *Coll. Math.* **5,** 198–204.

Schinzel, A. (1967). Reducibility and irreducibility of polynomials and covering systems of congruences. *Acta Arith.* **13,** 91–101.

Utz, W. (1961). A conjecture of Erdős concerning consecutive integers. *Amer. Math. Monthly.* **68,** 896–697.

Wirsing, E. (1967). Das asymptotische Verhalten von Summen über multiplikative Funktionen. *Acta Math. Acad. Sci. Hung.* **18,** 411–467.