SOME UNSOLVED PROBLEMS

Paul Erdős

On November 16, 1957, Assumption University of Windsor sponsored a symposium for mathematicians from Ontario, Michigan, and Indiana. The symposium gave occasion for an informal lecture in which I discussed various old and new questions on number theory, geometry and analysis. In the following list, I record these problems, with the addition of references and of a few further questions.

A. NUMBER THEORY

1. It is known [35, Vol. 1, Section 58] that \( \pi(2x) < 2\pi(x) \) for sufficiently large \( x \). Is it true that

\[
\pi(x + y) \leq \pi(x) + \pi(y) \ ?
\]

Ungár has verified the inequality for \( y \leq 41 \). Hardy and Littlewood [29, p. 69] have proved that

\[
\pi(x + y) - \pi(x) < cy/\log y.
\]

In the same paper, they discuss many interesting conjectures. They put

\[
\limsup_{x \to \infty} [\pi(x + y) - \pi(x)] = \rho(y),
\]

and they conjecture that \( \rho(y) > y/\log y \), and that perhaps \( \pi(y) - \rho(y) \to \infty \) as \( y \to \infty \).

Hardy and Littlewood deduce (2) by Brun's method. A very difficult conjecture, weaker than (1) but much stronger than (2), is that corresponding to each \( \varepsilon > 0 \) there exists a \( y_\varepsilon \) such that, for \( y > y_\varepsilon \),

\[
\pi(x + y) - \pi(y) < (1 + \varepsilon)y/\log y.
\]

It has not yet been disproved that \( \rho(y) = 1 \) for all \( y \). If \( \rho(y) > 1 \) for some \( y \), then \( \liminf (p_{n+1} - p_n) < \infty \).

2. About seventy years ago, Piltz [38] conjectured that, for each \( \varepsilon > 0 \),

\[
p_{n+1} - p_n = O(n^{\varepsilon}).
\]

Cramér conjectured [7, p. 24] that \( p_{n+1} - p_n = O((\log n)^2) \). If

\[
\limsup (p_{n+1} - p_n)/(\log n)^2 = 1,
\]

then, for each \( \varepsilon > 0 \), infinitely many of the intervals \([n, n + (1 - \varepsilon)(\log n)^2]\) contain no primes, but for \( n > n_\varepsilon \), there is a prime between \( n \) and \( n + (1 + \varepsilon)(\log n)^2 \). The Riemann hypothesis implies that

\[
p_{n+1} - p_n < n^{\varepsilon + 1/2}
\]

[35, Vol. 1, p. 338]. Thus the old conjecture that there is always a prime between two consecutive squares already goes beyond the Riemann hypothesis.
The first big achievement in this direction is due to Hoheisel [32], who showed
that \( p_{n+1} - p_n < n^{1-\delta} \). Ingham [34, p. 256] proved that \( 1 - \delta \) can be taken to be \( 5/8 \). In the opposite direction, I have proved [9, p. 124] that, for a certain \( c > 0 \) and for
infinitely many \( n \),

\[
p_{n+1} - p_n > \frac{c \log p_n \log \log p_n}{(\log \log p_n)^2}.
\]

Rankin [40] has shown that the factor \( \log \log \log \log p_n \) can be inserted in the numer-
ator; but this result seems to constitute the "natural boundary" for our method.

3. Let \( d_n = p_{n+1} - p_n \). Conjecture: The two sets of integers \( n \) for which
\( d_{n+1} \geq d_n \) and \( d_{n+1} \leq d_n \), respectively, have densities \( 1/2 \); and for infinitely many
\( n \), \( d_{n+1} = d_n \). It is only known (see [13]) that the set of integers \( n \) for which
\( d_{n+1} \geq (1 + \varepsilon)d_n \) has positive density, for some \( \varepsilon > 0 \). Does it happen
often that \( d_{n+2} > d_{n+1} > d_n \)? or that \( d_{n+2} < d_{n+1} < d_n \)? (If the answer is negative
in both cases, then the sequence \( \{d_{n+1} - d_n\} \) has ultimately alternating signs.) For
further questions and results on the difference of consecutive primes, see [24] and
[27].

4. The set of limit points of the sequence \( \{(p_{n+1} - p_n)/\log p_n\} \) has positive
measure [42, p. 94], but no finite limit point of the sequence is known. Is the se-
quence dense on the positive axis? Does it have a distribution function? For
further questions on this topic, see [42].

5. Chowla [6] has proved that there exist infinitely many triplets of primes in
arithmetic progression. Do there exist arbitrarily long arithmetic progressions of
primes? The longest of the known arithmetic progressions of primes has ten terms:
\( 199 + 210 h \) (\( h = 0, 1, \ldots, 9 \)). An old conjecture: There exist arbitrarily long arith-
metic progressions of consecutive primes.

6. A well-known theorem of Van der Waerden [50] (see also [39]) states that
there exists a function \( f(k) \) such that if the integers from 1 to \( f(k) \) are divided into
two sets, then at least one of the sets contains an arithmetic progression of \( k \) terms.
But all the known functions \( f(k) \) increase so rapidly that they do not even satisfy the
condition

\[
f(k) = k^{k^{\cdots^k}} \quad (k \text{ exponents}).
\]

Rado and I [23, p. 438] have shown that \( f(k) \geq \sqrt{k^{2^{k+1}}} \). This should be improved;
especially, a majorant for the least of the functions \( f(k) \) should be found.

7. Cramér [7, p. 45] showed that if the Riemann hypothesis is true, then, for
each \( \varepsilon > 0 \),

\[
\sum_{p_n < x} (p_{n+1} - p_n)^2 < c \varepsilon x(\log x)^{3+\varepsilon}.
\]

Perhaps the exponent \( 3 + \varepsilon \) could be reduced to \( 1 \); but the proof seems very difficult.
The following analogous conjecture [11, p. 440] is perhaps not quite hopeless: There
exists an absolute constant \( c \) such that if
are the \( \phi(n) \) integers less than \( n \) and relatively prime to \( n \), then

\[
\sum_{k=1}^{\phi(n)-1} (a_{k+1} - a_k)^2 < cn^2/\phi(n).
\]

8. The well-known conjecture on twin primes can be stated in the form

\[
\lim \inf d_n = 2. \quad \text{It is not known whether } \lim \inf d_n/\log n = 0. \quad \text{In this direction, I have proved [14] that } \lim \inf d_n/\log n < 1, \quad \text{and Rankin [41] has shown that the constant 1 can be replaced by } 57/59.
\]

I have shown [14] that

\[
\lim \sup \frac{\min(d_n, d_{n+1})}{\log n} = \infty;
\]

but I could not prove that

\[
\lim \sup \frac{\min(d_n, d_{n+1}, d_{n+2})}{\log n} = \infty,
\]

nor that

\[
\lim \inf \frac{\max(d_n, d_{n+1})}{\log n} < 1.
\]

For many further problems on this, see the forthcoming paper by Schinzel and Sierpiński in the Acta Arithmetica.

9. Twenty-five years ago, I conjectured that if \( f(n) = \pm 1 \) for \( n = 1, 2, \ldots \), and \( C \) is any constant, then there exist integers \( d \) and \( m \) such that

\[
\left| f(d) + f(2d) + \cdots + f(md) \right| > C
\]

(compare Van der Waerden's theorem in Problem 5). The conjecture clearly implies that if the function \( f(n) \) is also required to be multiplicative, then the sequence \( \{\sum_{k=1}^{n} f(k)\} \) is unbounded. For the special case where \( f(k) \) is the multiplicative function \( \lambda(k) \) equal to -1 whenever \( k \) is prime, we have

\[
\sum_{n=1}^{\infty} \frac{\lambda(n)}{n^s} = \frac{\zeta(2s)}{\zeta(s)}
\]

[35, Vol. 2, p. 618]; the series can not converge for \( s = \sigma + it \), if \( \sigma < 1/2 \). Therefore

\[
\sum_{n<x} \lambda(n) \neq O(x^{1/2-\varepsilon}).
\]

The Riemann hypothesis implies that
\[ \sum_{n \leq x} \lambda(n) = O(x^{1/2+\varepsilon}). \]

It has also been conjectured, by many mathematicians, that if \( f(n) \) is multiplicative and \( f(n) = \pm 1 \), then

\[ \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} f(k) \]

exists. For the case where \( f(k) = \lambda(k) \), the result follows from the prime number theorem. An example of Wintner shows that if we only assume that \( f \) is a complex-valued, multiplicative function with \( |f(k)| = 1 \), then the limit (3) need not exist. (See [51, p. 48], also [48].)

10. Let \( r_k(n) \) be the largest number of positive integers less than \( n \) which contains no arithmetic progression of \( k \) terms. The first publication on the function \( r_k(n) \) is due to Turán and myself [25]; we proved that \( r_3(2n) < n+1 \), for \( n > 7 \). We had been motivated by the fact that the inequality \( r_k(n) < n/2 \) would imply Van der Waerden's theorem (see Problem 6); but the problem itself seems to be much older (it seems likely that Schur gave it to Hildegard Ille, in the 1920's).

In [25], we stated our conjecture that \( \lim r_3(n)/n = 0 \), and Szekeres' conjecture that \( r_3[(3^k + 1)/2] = 2^k \). The latter conjecture is correct for \( k = 1, 2, 3 \); and we claimed that it holds for \( k = 4 \) (in other words, we claimed that \( r_3(41) = 16 \). We had deduced this from the relation \( r(20) = 8 \), which we had "proved" by trial and error. Recently, A. Mąkowski showed that \( r_3(20) = 9 \) and that \( r_3(18) = r_3(19) = 8 \) (private communication).

Behrend [2] proved that the limits \( c_k = \lim_{n \to \infty} r_k(n)/n \) exist, and that as \( k \to \infty \), either \( c_k \to 0 \) (in which case \( c_k = 0 \) for all \( k \)) or \( c_k \to 1 \).

Salem and Spencer [45], [46] disproved the Szekeres conjecture by showing that

\[ r_3(n) > n^{1-c/\log \log n}, \]

and Behrend [3] improved on this by showing that

\[ r_3(n) > n^{1-c/\sqrt{\log n}}. \]

Finally, Roth [44] proved that \( r_3(n) = o(n) \), more precisely, that \( r_3(n) < cn/\log \log n \). The true order of magnitude of \( r_3(n) \) and, more generally, of \( r_k(n) \), remains unknown.

11. Is it possible to choose \( k+2 \) integers \( a_1 \leq a_i \leq 2^k \) such that the \( 2^k \) possible sums \( \sum_{i=1}^{k} \varepsilon_i a_i \) (\( \varepsilon_i = 0 \) or \( \varepsilon_i = 1 \)) are all different? More generally, let \( h(x) \) be the maximum number of integers \( a_1 \leq a_i \leq x \) such that the corresponding sums are all different. Then \( h(x) > (\log x)/\log 2 \), as can be seen by choosing all the numbers \( 2^a, 2^1, \ldots \) that do not exceed \( x \). Moser and I have proved [19, p. 137] that

\[ h(x) < \frac{\log x}{\log 2} + (1 + \varepsilon) \frac{\log \log x}{2 \log 2}. \]

Perhaps \( h(x) = (\log x)/\log 2 + O(1) \).
12. Let \( A = \{a_i\} \) be a set of nonnegative integers such that each natural number \( n \) has a representation \( n = a_i + 2^j \), and let \( k(x) \) denote the number of \( a_i \) that are less than \( x \). It is known [36, Theorem 2] that there exists a set \( A \) for which
\[
k(x) < cx(\log \log x)/\log x.
\]
Can the factor \( \log \log x \) be dropped?

There exists a set \( \{a_i\} \) such that each natural number has a representation \( n = a_i + p_j \) (\( p_j \) prime), and such that
\[
k(x) < c(\log x)^2
\]
(see [18, p. 847]; this paper contains several further problems). Can this result be improved?

Hanani (oral communication) has raised the following question: Let \( \{a_i\} \) and \( \{b_j\} \) be two increasing sequences of natural numbers such that each \( n \) has a representation \( n = a_i + b_j \); is it true that
\[
\lim \sup \frac{k_a(x) k_b(x)}{x} > 1?
\]

13. Let \( \{a_i\} \) be an increasing sequence of natural numbers, and let \( f(n) \) denote the number of solutions of the equation \( n = a_i + a_j \). Turán and I [26, p. 215] conjectured that if \( f(n) > 0 \) for all \( n \), then \( \lim \sup f(n) = \infty \); and, more generally, that if \( a_k < ck^2 \) for all \( k \), then \( \lim \sup f(n) = \infty \). The best we have been able to prove, so far, is that \( \lim \sup f(n) \geq 2 \). For the literature on this and on related questions, see [22], also the comprehensive paper by Stöhr [47].

14. Each integer satisfies one of the following five congruences:

\[
x \equiv 0 \pmod{2}, \quad x \equiv 0 \pmod{3}, \quad x \equiv 1 \pmod{4}, \quad x \equiv 5 \pmod{8}, \quad x \equiv 7 \pmod{12}.
\]

If \( n_0 \) is a positive integer, does there exist an integer \( k \) and a system of \( k \) congruences

\[
x \equiv a_i \pmod{n_i} \quad (n_0 < n_1 < \cdots < n_k)
\]
such that each integer satisfies at least one of them?

Davenport and I have found such a system with \( n_1 = 3 \). Swift (oral communication) has found systems with \( n_1 = 4 \) and \( n_1 = 6 \), respectively; but no general methods are known. Does there exist a system in which all the \( n_i \) are odd? For the literature, see [15] and [17].

15. Let \( \{a_i\} \) be the increasing sequence of natural numbers of the form \( x^2 + y^2 \). Estimate \( a_{k+1} - a_k \). This seems to be an old problem. Bambah and Chowla [1] showed easily that \( a_{k+1} - a_k = O(a_k^{1/4}) \); but it has not yet been shown that the \( O \) can be replaced by \( o \). In the opposite direction, I have proved [16] that, for infinitely many \( k \),

\[
a_{k+1} - a_k > c(\log a_k)/\log \log a_k.
\]
B. GEOMETRY

16. Let \( S = \{p_1, \ldots, p_n\} \) be a finite point set in the plane; let \( N(S) \) denote the number of different values which occur among the distances \( p_ip_j \) \( (i \neq j) \), and \( f(n) \) the least value that \( N(S) \) can have. In [12], I proved that

\[
\frac{cn}{\log n} > f(n) > \sqrt[n]{n - 1} - 1;
\]

later, Moser [37] showed that \( f(n) > \frac{cn^2}{\sqrt{n}} \). It can be conjectured that, for \( n > n_c \), \( f(n) > n^{1-\varepsilon} \), and further, that there exists a point \( p_i \) such that more than \( n^{1-\varepsilon} \) different values occur among the distances \( p_ip_j \).

In [12], I conjectured that if the \( n \) points of \( S \) are the vertices of a convex \( n \)-gon, then at least \( \lceil n/2 \rceil \) of the distances are different; Moser [37] showed that at least \( \lceil (n+2)/3 \rceil \) of the distances are different. A stronger conjecture: Every convex \( n \)-gon has a vertex such that no three vertices are equidistant from it.

17. Borsuk [4] conjectured that in Euclidean \( n \)-space each set of diameter 1 is the union of \( n + 1 \) sets of diameter less than 1. For \( n = 2 \), the proof is easy. For \( n = 3 \), the first proof was given by Eggleston [8]. Simpler proofs have since been found by Grünbaum [28] and by Heppes [31]. For \( n > 3 \), the question remains open.

18. If \( S \) is a set of \( n \) points in the plane, of diameter 1, then the distance 1 can occur at most \( n \) times [33]. Vázsonyi (oral communication) conjectured that in 3-space the distance 1 can occur at most \( 2n - 2 \) times, and this was proved independently by Heppes [30], by B. Grünbaum and by Strasziewicz (see Math. Reviews soon). For \( n \)-space \( (n > 3) \), the problem remains open.

19. Conjecture: If \( S \) is a set of \( 2^k + 1 \) points in \( k \)-space, then three of the points determine an obtuse angle. For \( k = 2 \), this is trivial; for \( k = 3 \), an unpublished proof has been given by N. H. Kuiper and by A. H. Boerdijk.

Does there exist a set of six (or seven) points in 3-space such that all the plane angles determined by them are acute? For further problems of this type, see [12] and [21].

C. ANALYSIS

20. Let \( f(\theta) \) be a trigonometric polynomial whose roots are real and whose maximum absolute value on \([0, 2\pi]\) is 1. Conjecture:

\[
\int_0^{2\pi} |f(\theta)| \, d\theta \leq 4.
\]

This was proposed, with many similar problems, in [10].

21. Let \( |z_k| = 1 \) \((k = 1, 2, \cdots)\), \( f_n(z) = \prod_{k=1}^n (z - z_k) \). Conjecture: The sequence \( \{f_n(z)\} \) can not be uniformly bounded on \( |z| = 1 \).

Let the discrepancy of \( \{z_1, z_2, \ldots, z_n\} \) \((|z_k| = 1)\) be defined by the formula

\[
d(z_1, \ldots, z_n) = \max_{|a| = 1, \ |b| = 1} \left( N_n(a, b) - \frac{b - a}{2\pi n} \right),
\]
where $\overarc{ab}$ denotes the length of arc from $a$ to $b$, and $N_n(a, b)$ is the number of points $z_k$ ($k \leq n$) lying on this arc; clearly, the discrepancy is a measure of the greatest deviation from uniform distribution. Van der Corput conjectured (and Mrs. van Aardenne-Ehrenfest proved that), for any infinite sequence $\{z_k\}$ on $|z| = 1$, 
\[
\limsup_{n \to \infty} d(z_1, \ldots, z_n) = \infty.
\]
Later, Roth proved that
\[
\max_{1 \leq k \leq n} d(z_1, \ldots, z_n) > c \sqrt{\log n}.
\]

It has been conjectured by many that $c \log n$ is the correct minorant of the maximum (for references, see [43]).

Does there exist an infinite sequence $\{z_k\}$ ($|z_k| = 1$) such that, for each arc $(a, b)$ on $|z| = 1$, there exists a constant $c(a, b)$ with the property that, for all $n$,
\[
N_n(a, b) - \frac{b - a}{2\pi} n < c(a, b).
\]

22. If $P(z) = \sum_1^n a_k z^k$, where $|a_k| = 1$ for all $k$, then $\max_{|z|=1} |P(z)| > \sqrt{n}$, by Parseval’s Theorem. Does there exist a universal constant $c > 0$ such that, for $n \geq 2$, the maximum is at least $(1 + c)\sqrt{n}$? I have an unpublished proof that if
\[
P(\theta) = \sum_{k=1}^n \varepsilon_k \cos k\theta \quad (\varepsilon_k = \pm 1),
\]
then
\[
\max_{0 < \theta < 2\pi} |P(\theta)| > (1 + c)\sqrt{n/2}.
\]

23. Let $\{S_n\}$ be a sequence of sets none of which has a finite limit point. Does there exist an entire function $f(z)$ such that, for some sequence $\{k_n\}$, $f^{(k_n)}(z) = 0$ throughout $S_n$? (See [20].)

24. Does there exist an entire function $f$, not of the form $f(z) = a_n + a_n z$, such that the number $f(x)$ is rational or irrational according as $x$ is rational or irrational? More generally, if $A$ and $B$ are two denumerable, dense sets, does there exist an entire function which maps $A$ onto $B$?

25. For an entire function $f$, let
\[
M(r) = \max_{\theta} |f(re^{i\theta})|, \quad \mu(r) = \max_n |a_n|r^n,
\]
Then $\mu(r) < M(r)$ (with trivial exceptions). Conjecture: If $\lim_{r \to \infty} \mu(r)/M(r)$ exists, the limit is 0.

26. Does there exist, for each $n$, a polynomial
\[
f_n(z) = \sum_{k=1}^n \varepsilon_{nk} z^k \quad (\varepsilon_{nk} = \pm 1)
\]
such that, for all \( \theta \), \( c_1 \sqrt{n} < |f_n(e^{i\theta})| < c_2 \sqrt{n} \), where \( c_1 \) and \( c_2 \) are positive constants independent of \( \theta \) and of \( n \). Cluny has some (unpublished) partial results on this.

27. I call attention to the following beautiful problem of Turán [49, p. 27] (see also Cassels [5]). Let \( z_i = 1, |z_i| = 1 \) (i = 2, 3, \ldots, n); and let \( s_k = \sum_{i=1}^{n} z_i^k \). Does there exist a constant \( c \), independent of \( n, k \) and \( \{z_i\} \), such that

\[
\max_{1 < k < n} |s_k| > c?
\]

The best result is due to Bruijn (see the Chinese edition of Turán's book):

\[
\max_{1 < k < n} |s_k| > c(\log \log n)/\log n.
\]

28. There exists a sequence \( \{s_n\} \) of 0's, 1's and 2's such that no two adjacent blocks from \( \{s_n\} \) are the same. Presumably, the first (unpublished) proof of this was obtained by Rose Peltesohn and J. W. Sutherland.

Let \( N(k) \) be the least number \( N \) with the property that each sequence \( \{s_n\}_{n=1}^{N} \) of numbers taken from the set \( \{1, 2, \ldots, k\} \) contains two adjacent blocks such that each is a rearrangement of the other. My earliest conjecture, that \( N(k) = 2^k - 1 \), has been disproved by Bruijn and myself. It is not even known whether \( N(4) < \infty \).

In conclusion, I express my thanks for the warm hospitality with which Assumption University received its guests.

REFERENCES


17. \textemdash, \textit{On a problem concerning congruence systems}, Mat. Lapok 3 (1952), 122-128. (Hungarian, Russian and English summaries.)


20. \textemdash, \textit{Remarks on a paper by Kőváry}, Mat. Lapok 7 (1956), 214-217. (Hungarian; English and Russian summaries.)

21. \textemdash, \textit{On some problems in geometry}, Mat. Lapok 8 (1957), 81-92. (Hungarian; English and Russian Summaries.)


38. A. Piltz, Über die Häufigkeit von Primzahlen in arithmetischen Progressionen ..., Habilitationsschrift, Jena, 1884.
46. ———, On sets which do not contain a given number of terms in arithmetical progression, Nieuw Arch. Wiskunde (2) 23 (1952), 133-143.
50. B. L. Van der Waerden, Beweis einer Baudet'schen Vermutung, Nieuw Arch. Wiskunde (2) 15 (1928), 212-216.
51. A. Wintner, The theory of measure in arithmetical semi-groups, Baltimore, 1944.