ტფილისის მათემატიკური ინსტიტუტის შრომები, ტ. III, 1938
ТРУДЫ ТБИЛИССКОГО МАТЕМАТИЧЕСКОГО ИНСТИТУТА, Т. III, 1938
TRAVAUX DE L'INSTITUT MATHÉMATIQUE DE TBILISSI, T. III, 1938

ON THE ASYMPTOTIC DENSITY OF THE SUM OF TWO SEQUENCES ONE OF WHICH FORMS A BASIS FOR THE INTEGERS. II.

BY PAUL ERDÖS (Manchester)

Let $a_1, a_2, \ldots$ be any sequence of positive steadily increasing integers, and suppose there are $x = f(n)$ of them not exceeding a given number n, so that

$$a_x \leqq n < a_{x+1}.$$

The Schnirelmann density δ of the sequence is defined as the lower bound of the numbers $f(n)/n$; $n = 1, 2, \ldots$. Thus, if $a_1 \neq 1$, $\delta = 0$. Clearly $f(n) \geqq \delta n$.

The asymptotic density δ_a of the sequence is defined as $\lim_{n \to \infty} \inf f(n)/n$.

Suppose also that the steadily increasing set of positive integers

$$A_0 = 0,\ A_1,\ A_2,\ \ldots$$

form a basis of order l of the positive integers. This means that every positive integer can be expressed as the sum of at most l of the A's.

In a previous paper[1], I proved the following

Theorem. *If δ' is the Schnirelmann density of the sequence $a + A$, i. e. of the integers which can be expressed as the sum of an a and an A, then*

$$\delta' \geqq \delta + \frac{\delta(1-\delta)}{2l}.$$

In § 1 of the present paper I prove the following

Theorem. *Let $a_1 < a_2 < \ldots$ be a sequence of positive integers of asymptotic density δ_a. Let $A_0 = 0$, $A_1 < A_2 < \ldots$ be a sequence of positive integers such that for every $\varepsilon > 0$, an M exists so that every integer $m \geqq M$ is the sum of at most l positive and negative A's, where the absolute value of the negative A's is less than εm. Then, if δ'_a is the asymptotic density of the sequence $a + A$,*

$$\delta'_a \geqq \delta_a + \frac{\delta_a(1-\delta_a)}{2l}.$$

[1] «On the arithmetical density of the sum of two sequences one of which forms a basis for the integers». Acta Arithmetica 1 (1936), 197—200. I shall refer to this paper as I.

In § 2, by aid of this result, I prove that *every large integer is the sum of two primes and a bounded number of squares of primes*. This may be contrasted with the result proved in a previous paper [2], that every integer is the sum of a bounded number of positive and negative squares of primes. It was conjectured that every large integer is the sum of a bounded number of positive squares of primes [3].

In § 3, I consider results analogous to the following theorem of Khintchine: Let $a_0=0<a_1<\dots$, $b_0=0<b_1<\dots$ be two sequences of Schnirelmann density $\delta \leqq \frac{1}{2}$, then the Schnirelmann density of the sequence $a+b$ is $\geqq 2\delta$.

Dr Heilbronn and I conjectured that if $a_0=0$, $a_1=1<a_2<\dots$, and $b_0=0$, $b_1=1<b_2<\dots$, are two sequences of asymptotic density $\delta_a \leqq \frac{1}{2}$, then the asymptotic density of the sequence $a+b$ is $\geqq \frac{3}{2}\delta_a$. I prove this conjecture in the special case when the two sequences a and b are identical, i. e. the following

Theorem. *Let $a_0=0$, $a_1=1<a_2<\dots$ be a sequence of integers of asymptotic density $\delta_a \leqq \frac{1}{2}$, then for the asymptotic density δ'_a of the sequence a_i+a_j,*

$$\delta'_a \geqq \frac{3}{2}\delta_a.$$

On the other hand, it is easy to see that if $\delta_a > \frac{1}{2}$, every sufficiently large integer is of the form a_i+a_j i. e. the asymptotic density of the sequence a_i+a_j is 1.

It is easy to see that this theorem is best possible. For let $a_0, a_1, a_2, \dots$ be all the integers $\equiv 0$, or 1 (mod 4). The asymptotic density of this sequence is 1/2. The sequence a_i+a_j consists of the integers $\equiv 0,1$ or 2 (mod 4) and its asymptotic density is 3/4. This example is due to Dr Heilbronn.

§ 1

The argument of this chapter is very similar to that of I. As there, we prove our theorem as a particular case of a more general one. Let n be sufficiently large and let the positive integers $\leqq n$ not included among the a's be denoted by $b_1, b_2, \dots, b_y$.

[2] «On the easier Waring Problem for powers of primes I». Proceedings of the Cambridge Philosophical Society, Vol. XXXIII. Part 1. January 1937, 6—12.

[3] Since this paper was written, this conjecture has been proved by Vinogradoff «Einige allgemeine Primzahlsätze». Travaux de l'Institut Mathématique de Tbilissi III (1938), 35—67.

Put

$$E = b_1 + b_2 + \ldots + b_y - \frac{1}{2} y(y+1),$$

so that $E \geqq 0$, since $b_1 \geqq 1$, $b_2 \geqq 2$ etc. Then I prove the existence of at least $x + \frac{E}{ln} - \frac{M}{l} - 2\varepsilon n$ integers $\leqq n$ of the form $a + A$, where in fact only $A = 0$ and a single other A need be used. This is deduced from the result that at least $\frac{E}{ln} - \frac{M}{l} - 2\varepsilon n$ of the b's can be represented in the form $a + A$ where in fact only a single A is used.

We require two lemmas.

Lemma 1. *If M is any given integer with $0 < M < \frac{E}{n}$, an integer $I > M$ exists such that there are at least $\frac{E}{n} - M$ of the b's $\leqq n$ in the set $a_1 + I$, $a_2 + I$, ...*

For the equation

$$a + v = b$$

has at least $E - nM$ solutions in the variables $v > M$, a, b. Thus, for given $b = b_r$ there are at least $b_r - r$ of the a's not less than b_r and hence at least $b_r - r - M$ of the a's not less than $b_r - M$. We find from each of these a's a solution v and summing for $r = 1, 2, \ldots, y$, the total number of solutions is not less than

$$\sum_{r=1}^{y} (b_r - r) - yM > E - nM.$$

But there are at most $n - M$ possible values of v, namely, $M+1$, $M+2$, ..., n, and so, for at least one value of v, say I, there are not less than

$$\frac{E - Mn}{n - M} > \frac{E}{n} - M$$

of the b's in $a + I$. This proves the lemma.

Lemma 2. *If ξ is the number of b's $\leqq n$ in the set $a + U$ where U is any given integer, and η is the number of b's in $a - U$, then $\eta \leqq \xi + U$.*

Let us denote by $a_1 < a_2 < \ldots < a_z$ the a's not exceeding $n - U$. Evidently $z \geqq x - U$. Thus the number of a's in the sequence $a_1 + U$, $a_2 + U$, ..., $a_z + U$ is not less than $z - \xi \geqq x - \xi - U$, hence the number of a's in $a_i - U$ is also not less than $x - \xi - U$. Thus the number of b's in $a_i - U$ does not exceed $\xi + U$, which proves the lemma.

Now we proceed to prove our main theorem. We express I as the sum and difference of exactly l of the A's say

$$I=\sum_{i=1}^{l}\varepsilon_i A_i,\ \varepsilon_1=\varepsilon_2=\ldots=\varepsilon_r=1,\ \ \varepsilon_{r+1}=\varepsilon_{r+2}=\ldots=\varepsilon_l=-1,$$

by including a sufficient number of A_0's among the A's if need be and where A_1 need not denote the first, A_2 the second etc. of the A's.

Denote by μ_s the number of b's $\leqq n$ in the set $a+\varepsilon_s A_s$; $s=1, 2, \ldots, l$. I prove now that

$$\mu_1+\mu_2+\ldots\mu_l+A_{r+1}+A_{r+2}+\ldots+A_l\geqq\frac{E}{n}-M.$$

For in the set of integers given by $a+A_1+A_2$ there are at most $\mu_1+\mu_2$ of the b's. Thus the set $a+A_1$ contains μ_1 of the b's together with some a's. When we add A_2 to the numbers of the set $a+A_1$, the μ_1 b's, give at most μ_1 b's, while the a's give at most μ_2 b's. Now take the set $a+A_1+A_2+A_3$. This contains at most $\mu_1+\mu_2+\mu_3$ of the b's by precisely the same arguments applied to the sum of $a+A_1+A_2$ and A_3. Similarly the set $a+A_1+A_2+\ldots+A_r$ contains at most $\mu_1+\mu_2+\ldots+\mu_r$ of the b's. We now assert that the set $a+A_1+A_2+\ldots+A_r-A_{r+1}$ contains at most $\mu_1+\mu_2+\ldots+\mu_r+\mu_{r+1}+A_{r+1}$ of the b's. For if we subtract A_{r+1} from the members of the set $a+A_1+A_2+$ $+\ldots+A_r$, the $\mu_1+\mu_2+\ldots+\mu_r$ b's give at most $\mu_1+\mu_2+\ldots+\mu_r$ b's while the a's give at most μ_{r+1} of them. Also the members of the set $a+A_1+A_2+\ldots+A_r$ exceeding n give at most A_{r+1} b's. By the same argument, the set $a+A_1+$ $+A_2+\ldots+A_r-A_{r+1}-\ldots-A_l$ i. e. the set $a+I$ contains at most $\mu_1+\mu_2+$ $+\ldots+\mu_l+A_{r+1}+A_{r+2}+\ldots+A_l$ of the b's. But by lemma 1 the set $a+I$ contains at least $\frac{E}{n}-M$, of the b's and hence the result.

$A_{r+1}, A_{r+2}, \ldots, A_l$ are all less than εn, thus we have

$$\mu_1+\mu_2+\ldots+\mu_l>\frac{E}{n}-M-l\,\varepsilon n.$$

Hence one of the μ's, say, $\mu_k\geqq\frac{E}{ln}-\frac{M}{l}-\varepsilon n$, and so if $k\leqq r$, or from lemma 2. by taking $U=A_k$ if $k>r$, the number of b's in $a+A_k$ is not less than $\frac{E}{ln}-2\varepsilon n-\frac{M}{l}$.

We may suppose without loss of generality that the a's have asymptotic density δ_a, say δ, with $\delta<1$. We have $f(b_\rho)\geqq(\delta-\eta)\,b_\rho$, every $\eta>0$ if $b_\rho>N=N(\eta)$; hence

$$b_\rho-\rho=f(b_\rho)\geqq(\delta-\eta)\,b_\rho,\quad b_\rho>\frac{\rho}{1-\delta+\eta},$$

and therefore

$$E=b_1+b_2+\ldots+b_y-\frac{y(y+1)}{2}\geqq\frac{1+2+\ldots+y}{1-\delta+\eta}-\frac{y(y+1)}{2}-N_1(\eta)$$

$$\geqq\frac{\delta-\eta}{2(1-\delta+\eta)}y(y+1)-N_1$$

for sufficiently small η.

Hence for the number T of integers not exceeding n in the set $a+A$ we have

$$T\geqq x+\frac{\delta-\eta}{2(1-\delta+\eta)}\frac{y^2}{nl}-\frac{M}{l}-2\,\varepsilon n-N_1.$$

Write

$$x+\frac{\delta-\eta}{2(1-\delta+\eta)}\frac{y^2}{nl}=\varphi(x)\qquad(y=n-x).$$

For

$$x\geqq(\delta-\eta)\,n,$$

$$\varphi'(x)=1-\frac{\delta-\eta}{2(1-\delta+\eta)}\frac{2(n-x)}{nl}\geqq 1-\frac{\delta-\eta}{l}>0$$

i. e.

$$T\geqq\varphi(x)-\frac{M}{l}-2\,\varepsilon n-N_1\geqq\varphi[(\delta-\eta)\,n]-\frac{M}{l}-2\,\varepsilon n-N_1$$

$$=(\delta-\eta)\,n+\frac{\delta-\eta}{2(1-\delta+\eta)}\frac{(1-\delta+\eta)^2n^2}{nl}-\frac{M}{l}-2\,\varepsilon n-N_1.$$

Hence

$$T\geqq n\left(\delta-\eta-2\varepsilon+\frac{(\delta-\eta)(1-\delta+\eta)}{2l}\right)-\frac{M}{l}-N_1.$$

This proves the inequality

$$\delta'\geqq\delta+\frac{\delta(1-\delta)}{2l}$$

and establishes the result.

§ 2

Let p, p_1, …, q, q_1, … denote primes; k a positive integer. Romanoff[4] proved that the density of the integers of each of the forms $p+k^2$ and $p+2^k$ is positive. By his method, I can prove that the density of integers of the form $p+q^2$ is positive. I have, however, proved in my paper[5] that every

[4] «Über einige Sätze der additiven Zahlentheorie». Mathematische Annalen, Band 109 (1934), 668—678.

[5] Loc. cit.

integer m is the sum of a bounded namber of positive and negative squares of primes. The proof shows that the primes in the representation of m may all be taken less than m. I prove now that the primes whose squares have a negative sign may be taken less than $m^{\frac{49}{50}}$, by aid of a result of Tchudakoff[6], namely, that for sufficiently large n the interval n, $n+n^{\frac{3}{4}+\varepsilon}$ contains at least one prime.

For suppose m is sufficiently large, and p_1 is the greatest prime not exceeding $m^{\frac{1}{2}}$; then from Tchudakoff's result, we obtain

$$m-p_1^2=(m^{\frac{1}{2}}-p_1)\,(m^{\frac{1}{2}}+p_1)<m^{\frac{7}{8}+\varepsilon}.$$

If now p_2 is the greatest prime not exceeding $m-p_1^2$, then $m-p_1^2-p_2^2<m$, and similarly

$$m-\sum_{i=1}^{6}p_i^2<m^{\left(\frac{7}{8}+\varepsilon\right)^6}<m^{\frac{49}{100}}.$$

Thus from the representation of the left hand side, a constant exists such that every sufficiently large integer is the sum of l positive and negative squares of primes where the negative squares may be supposed to be less than $m^{\frac{49}{50}}$.

Let the asymptotic density of this sequence $p+q_1^2$ be $\delta^{(1)}$; then by § 1 the asymptotic density of the sequence $p+q_1^2+q_2^2$ is not less than

$$\delta^{(1)}+\frac{\delta^{(1)}(1-\delta^{(1)})}{2l}=\delta^{(2)}.$$

In the same way, the asymptotic density of the sequence $p+q_1^2+q_2^2+q_3^2$ is not less than $\delta^{(2)}+\dfrac{\delta^{(2)}(1-\delta^{(2)})}{2l}=\delta^{(3)}$.

Hence a constant c exists such that asymptotic density of the integers of the form $p+q_1^2+q_2^2+\ldots+q_c^2$ is greater than $1/2$. From this it follows immediately that every sufficiently large integer is of the form $p_1+p_2+\sum_{i=1}^{2c}q_i^2$.

§ 3

Let $a_1=1<a_2<\ldots<a_x\leqq n<a_{x+1}\ldots$ be a sequence of asymptotic density $\delta_a=\delta$, and let ε be an arbitrary number. Let m be the greatest integer such that $f(m)\leqq(\delta-\varepsilon)m$ but for $y>m$, $f(y)>(\delta-\varepsilon)y$. Then $m+1$ is an a, for if not $f(m+1)=f(m)\leqq(\delta-\varepsilon)m<(\delta-\varepsilon)(m+1)$. It is easy to see that the Schnirelmann density of the positive members of the sequence a_i-m is not

[6] «On the difference between two neighbouring prime numbers». Recueil Mathématique 1 (1936), 799-813.

less than $\delta-\varepsilon$. Hence from a result of Khintchine, it follows that the Schnirelmann density of the sequence $(a_i-m)+(a_j-m)$ i. e. the density of the sequence $\{a_i-m,\ a_i+a_j-2m\}$ the members of which are given by the sequences $a_i-m,\ a_i+a_j-2m$ is not less than $2(\delta-\varepsilon)$, i. e. for sufficiently large n the number of integers not exceeding n of the sequence $\{a_i+m,\ a_i+a_j\}$ is not less than $2(\delta-\varepsilon)n-2m$. Let now n be sufficiently large and denote by $a_{i_1}<a_{i_2}<\ldots<a_{i_\nu}$ the integers not exceeding $n-m$ which $a_{i_r}+m$ does not occur in the sequence a_i+a_j. Since $m+1$ is an a_i, it follows that

$$a_{i_r}+m \neq a_{i_{r-1}}+m+1$$

i. e.

$$a_{i_{r-1}}+1=a_{i_{r-1}}+a_1$$

is not an a_i. Thus for the number of τ integers of the sequence $\{a_i, a_i+a_j\}$ not exceeding n we have the inequality

$$T \geqq \max[x+\Gamma,\ (2\delta-2\varepsilon)n-2m-\Gamma] \geqq \max[(\delta-\varepsilon)n+\Gamma,(2\delta-2\varepsilon)n$$
$$-2m-\Gamma] \geqq \frac{3}{2}(\delta-\varepsilon)n-2m.$$

This means that the asymptotic density of the sequence a_i+a_j is not less than $\frac{3}{2}\delta$ and proves the result.

University Manchester.

(Received 27 July, 1937.)

П. ЭРДЕШ

АСИМПТОТИЧЕСКАЯ ПЛОТНОСТЬ СУММЫ ДВУХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ, СОСТАВЛЯЮЩИХ БАЗИС ЦЕЛЫХ ЧИСЕЛ. II.

(Резюме)

Пусть асимптотическая плотность δ_a последовательности (a) целых чисел

$$a_1<a_2<\ldots$$

определена так:

$$\delta_a=\liminf_{n\to\infty}\frac{1}{n}\sum_{a_m\leqq n} 1.$$

Путем элементарных метрических соображений автор доказывает следующую теорему:

Если δ_a *асимптотическая плотность последовательности* (a); $A_0=0$, $A_1<A_2<\dots$ *последовательность* (A) *целых чисел; к любому положительному* ε *существует такое* M, *что всякое целое* $m\geqq M$ *представимо в виде суммы* l *слагаемых вида* $\pm A_j$, *причем модуль каждого отрицательного числа* $-A_j$ *меньше чем* εm; δ'_a *асимптотическая плотность последовательности* $(a+A)$, (т. е. совокупности чисел вида a_i+A_j)—

То

$$\delta'_a\geqq\delta_a+\frac{\delta_a(1-\delta_a)}{2l}.$$

Отсюда автор выводит, пользуясь одним из своих прежних результатов и оценкой

$$p_{n+1}-p_n=O\left(p_n^{\frac{3}{4}+\varepsilon}\right)$$

Чудакова: *Всякое достаточно большое целое представимо в виде суммы двух простых и ограниченного числа квадратов простых.*

(Этот результат превзойден теоремой 5 работы И. М. Виноградова «Некоторые общие теоремы, относящиеся к теории простых чисел» (см. стр. 29 этого тома Трудов), из которой следует, что всякое достаточно большое целое представимо в виде суммы не более девяти квадратов простых.)

Затем автор доказывает, опять элементарным путем:

Если $a_0=0$, $a_1=1<a_2<\dots$ *последовательность* (a) *целых чисел асимптотической плотности* $\delta_a\leqq\frac{1}{2}$, *то асимптотическая плотность последовательности* $(a+a)$ (т. е. совокупности чисел вида a_i+a_j) *по крайней мере равна*

$$\frac{3}{2}\delta_a.$$

Этой теоремы нельзя улучшить, как показывает пример

$$a_0=0,\ \ a_1=1,\ \ a_2=4,\ \ a_3=5,\ \ a_4=8,\ \ a_5=9,\ \ \dots .$$