ACTA ARITHMETICA
VII (1962)

On the representation of large integers as sums of distinct summands taken from a fixed set

by

P. ERDÖS (Haifa)

Let A be a sequence of integers $a_1 < a_2 < \dots$

$$A(n) = \sum_{a_i \leqslant n} 1 ,$$

i.e. $A(n)$ denotes the number of a's not exceeding n.

Some time ago I conjectured that if $(u, v) = 1$ then every sufficiently large integer is the sum of distinct summands of the form $u^k v^l$. Recently Birch [1] has proved this conjecture, his proof being elementary but ingenious and difficult.

Further I conjectured that if the sequence A satisfies $a_{k+1}/a_k \to 1$ and is such that every arithmetic progression contains at least one integer which is the sum of distinct a's, then every sufficiently large integer is the sum of distinct a's. If we further assume that $A(n) > n^{1-c_1}$ ($c_1, c_2, \dots$ denote positive absolute constants), then I have proved my conjecture, but this result has seemed of little interest since I have hoped that my conjecture is true.

Recently, however, Cassels [2] has proved the following theorems:

1. *Assume that*

$$\lim_{n\to\infty} \big(A(2n) - A(n)\big)/\log\log n = \infty$$

and that for every real θ, $0 < \theta < 1$

$$\sum_{k=1}^{\infty} \|a_k\theta\|^2 = \infty, \qquad \|a\| = \min_{-\infty<n<\infty} |a-n| .$$

Then every sufficiently large number is the sum of distinct a*'s.*

2. *For every* $\varepsilon > 0$ *and* $\eta > 0$ *there exists a sequence* A *containing infinitely many terms in every arithmetic progression and satisfying*

$$a_{n+1} - a_n = o(a_n^{1/2+\eta})$$

so that the number of integers $\leqslant x$ which are the sum of distinct a's is $< \varepsilon x$ for $x > x_0$.

It is easy to see that the first theorem of Cassels contains Birch's result. The ingenious proof of Cassels is analytic and uses the method of Hardy Littlewood.

The second theorem of Cassels clearly shows that my conjecture is wrong, but then my old result is perhaps not entirely without interest. In fact, I have succeeded in strengthening it somewhat. In this note I am going to prove the following:

THEOREM. *Let C be a sufficiently large integer, and $a_1 < a_2 < \dots$ an infinite sequence of integers satisfying*

$$(0)\quad A(x) > Cx^{(\sqrt{5}-1)/2} \quad \text{for } x > x_0, \qquad \text{or} \qquad a_r < \left(\frac{r}{C}\right)^{2/(1+\sqrt{5}-1)} \quad \text{for } r > r_0.$$

Assume further that every arithmetic progression contains at least one integer which is the sum of distinct a's. Then every sufficiently large integer is the sum of distinct a's.

It would be interesting to know, especially in view of the second theorem of Cassels, whether the exponent in (0) can be improved. I have not succeeded in doing this, but perhaps an improvement of my method will give the Theorem if (0) is replaced by $A(x) > x^{1/2+\varepsilon}$ for every $\varepsilon > 0$ if $x > x_0(\varepsilon)$. Perhaps the Theorem remains true if we only assume $A(x) > Cx^{1/2}$, but a simple argument shows that $A(x) > Cx^{1/2}$ is not sufficient if $C < \sqrt{2}$. In fact, the following simple result holds: Let $a_1 < a_2 < \dots$ Assume $a_k < \frac{k^2+ck}{2}$ where c is an absolute constant. Then for all sufficiently large k, $a_k < a_1 + a_2 + \dots + a_{k-1}$. It is easy to see that this result is the best possible in the following sense: Let β_k tend to infinity arbitrarily slowly with k. Then there exists a sequence $a_1 < a_2 < \dots$ satisfying $a_k < \frac{k^2+\beta_k k}{2}$ for which $\limsup_{k=\infty}\left(a_k - \sum_{i=1}^{k-1} a_i\right) = \infty$. This of course implies that there are infinitely many integers which are not sums of distinct a's. We leave the simple proofs of these statements to the reader.

First we prove three lemmas.

LEMMA 1. *Let n be sufficiently large, $Z > 10n^{1/2}$ and let $\frac{1}{2}n < b_1 < \dots < b_Z < n$ be any Z integers. Denote by $f(m)$ the number of solutions of $m = b_i + b_j$, $i \neq j$ ($b_i + b_j$ and $b_j + b_i$ are not counted as distinct solutions). Then there exists an integer k, $1 \leqslant k \leqslant \frac{\log n}{2\log 2}$ satisfying*

$$(1)\qquad f(u) > \frac{Z}{(k+1)^2 2^k}, \qquad f(v) > \frac{Z}{(k+1)^2 2^k}, \qquad v - u \leqslant \frac{10n}{Z 2^k}.$$

The number of sums $b_i + b_j$, $i \neq j$, is clearly equal to $\binom{Z}{2} > \frac{1}{3}Z^2$, also $n < b_i + b_j < 2n$. Thus

$$\sum_{n+1}^{2n} f(m) > \tfrac{1}{3}Z^2.$$

Hence there clearly are two integers c and d satisfying

$$(2)\qquad d - c = \left[\frac{10n}{Z}\right], \qquad \sum_{c}^{d} f(m) > \frac{10(Z-10)}{3}.$$

Consider the m's in (c, d) for which

$$(3)\qquad \frac{Z}{(k+1)^2 2^k} < f(m) \leqslant \frac{Z}{k^2 2^{k-1}} \qquad \left(k = 1, 2, \dots, \left[\frac{\log n}{2\log 2}\right]\right).$$

If there are more than 2^k integers satisfying (3) for some k, then two of them, say u and v, satisfy $v - u \leqslant 10n/2^k Z$ and (1) is satisfied. Thus to complete the proof of our Lemma it suffices to show that for some k there must be more than 2^k integers satisfying (3). Assume that this is false. Then since $f(m) < Z$ we obtain

$$(4)\qquad \sum_{c}^{d} f(m) \leqslant \sum_{k=1}^{l} 2^k \frac{Z}{k^2 2^{k-1}} + \sum_{c}^{d}{}' f(m) < Z\frac{\pi^2}{3} + \sum_{c}^{d}{}' f(m)$$

where $l = \left[\frac{\log n}{2\log 2}\right]$ and $\sum_{c}^{d}{}' f(m)$ is extended over those m in (c, d) for which $f(m) \leqslant \frac{Z}{(l+1)^2 2^l}$. Thus since $Z > 10n^{1/2}$

$$(4')\qquad \sum_{c}^{d}{}' f(m) \leqslant (d - c + 1)\frac{Z}{(l+1)^2 2^l} < 20\frac{n}{Z}\cdot\frac{10Z}{n^{1/2}(\log n)^2} = o(Z).$$

From (4) and (4') we have

$$\sum_{c}^{d} f(m) < \tfrac{1}{3}Z\pi^2 + o(Z) < \tfrac{10}{3}(Z-10) \qquad \text{for sufficiently large } Z,$$

which contradicts (2), and thus our Lemma is proved.

LEMMA 2. *Let G be a group of n elements and $a_1, a_2, \dots$ a finite or infinite sequence of elements of G. Suppose that there are k distinct elements $b_1, \dots, b_k$ of G which can be represented in the form $\prod a_i^{\varepsilon_i}$, $\varepsilon_i = 0$ or 1 (the product is always finite). Then there exist k or fewer a's, $a_{i_1}, \dots, a_{i_r}$, $r \leqslant k$, so that each of the b's can be represented in the form $\prod_{j=1}^{r} a_{i_j}^{\varepsilon_j}$.*

First of all we can assume that the number of a's is at least k, for otherwise our Lemma is trivial.

If the unit element of G cannot be represented in the form $\prod a_i^{\varepsilon_i}$, then all the elements $a_1, a_1 \cdot a_2, \ldots, a_1 \cdot a_2 \cdot \ldots \cdot a_k$ are distinct and our Lemma is proved. Thus we can assume that the unit can be represented in the for $m \prod a_i^{\varepsilon_i}$ and let $a_{i_1} \ldots a_{i_s}$ be its shortest representation $(s < k)$. There are at least s distinct elements of the form $\prod_{j=1}^{s} a_{i_j}^{\varepsilon_j}$ since the s elements $a_{i_1}, a_{i_1} \cdot a_{i_2}, \ldots, a_{i_1} \cdot \ldots \cdot a_{i_s}$ are all distinct.

If all the b's are of the form $\prod_{j=1}^{s} a_{i_j}^{\varepsilon_j}$, our proof is complete. If not, there exists an $a_{i_{s+1}}$ so that $\prod_{j=1}^{s+1} a_{i_j}^{\varepsilon_j}$ contains at least one b which is not of the form $\prod_{j=1}^{s} a_{i_j}^{\varepsilon_j}$ (for otherwise all the b's would be of the form $\prod_{j=1}^{s} a_{i_j}^{\varepsilon_j}$). If all the b's are of the form $\prod_{j=1}^{s+1} a_{i_j}^{\varepsilon_j}$, our proof is complete; otherwise, we can find an $a_{i_{s+2}}$ such that $\prod_{j=1}^{s+2} a_{i_j}^{\varepsilon_j}$ contains at least one b not of the form $\prod_{j=1}^{s+1} a_{i_j}^{\varepsilon_j}$. Continuing in this way we finally obtain $a_{i_1}, \ldots, a_{i_s}, a_{i_{s+1}}, \ldots, a_{i_{s+t}}$, so that every b is of the form $\prod_{j=1}^{s+t} a_{i_j}^{\varepsilon_j}$ and that each $a_{i_{s+r}}$, $1 \leqslant r \leqslant t$ gives at least one new $\underline{b}$, or $k \geqslant s+t$ and the proof of Lemma 2 is complete.

LEMMA 3. *Let $x_1, x_2, \ldots, x_r$ be any r integers. Then $\sum_{i=1}^{r} \varepsilon_i x_i \equiv 0 \pmod r$ is solvable in numbers $\varepsilon_i = 0$ or 1.*

Lemma 3 is well known. If the sums $x_1, x_1+x_2, \ldots, x_1+\ldots+x_r$ are all incongruent mod r, one of the sums is 0 and there is nothing to prove. If $x_1+\ldots+x_{k_1} \equiv x_1+\ldots+x_{k_2} \pmod r$, $k_1 < k_2 \leqslant r$, then $x_{k_1+1}+\ldots+x_{k_2} \equiv 0 \pmod r$, which proves the Lemma.

Now we can prove our Theorem. Put $a = \frac{1}{2}(\sqrt{5}-1)$ and

$$\limsup_{x=\infty} \frac{\log A(x)}{\log x} = \beta, \quad a \leqslant \beta \leqslant 1 \quad \text{by (0)}. \tag{5}$$

Assume first $\beta > a$. Let $\varepsilon = \varepsilon(a, \beta)$ be sufficiently small and choose sufficiently large n, so that

$$A(n) - A\left(\frac{n}{2}\right) > n^{\beta-\varepsilon}. \tag{6}$$

By (5) such n clearly exist. If $\beta = a$ we distinguish two cases. If

$$\limsup_{x=\infty} \frac{A(x)}{x^a} = D < \infty, \quad D \geqslant C, \tag{7}$$

we choose sufficiently large n, so that

$$A(n) - A\left(\frac{n}{2}\right) > \frac{D}{4} n^a. \tag{8}$$

By (7) such n clearly exist. If

$$\limsup_{x=\infty} \frac{A(x)}{x^a} = \infty \tag{9}$$

we choose n sufficiently large and such that

$$\frac{A(n)}{n^a} > \frac{A(m)}{m^a}, \quad 1 \leqslant m < n. \tag{10}$$

It follows from (9) that (10) can be satisfied for arbitrarily large n. Put

$$\left[\frac{1}{2}\left(A(n) - A\left(\frac{n}{2}\right)\right)\right] = Z \tag{11}$$

and apply Lemma 1 to the Z largest a's, $a_{l+1}, \ldots, a_{l+Z}$, in the interval $(n/2, n)$. From (1) it follows that there are two integers $n < u < v < 2n$ such that for a certain $k \leqslant \frac{\log n}{2 \log 2}$ the equations

$$a_i + a_j = u, \quad a_{i'} + a_{j'} = v, \quad T = v - u \leqslant \frac{10n}{Z 2^k}, \quad l+1 \leqslant i, j, i', j' \leqslant l+Z, \tag{12}$$

both have at least $[Z/(k+1)^2 2^{k+1}]$ solutions where each $\underline{a}$ occurs as a summand in at most one of the equations (12). To prove (12) observe that Lemma 1 implies that equations (12) both have more than $Z/(k+1)^2 2^k$ solutions. If an a_k occurs as a summand in both $a_i + a_j = u$ and $a_{i'} + a_{j'} = v$, we only count it as a solution of one of the equations, and we can clearly arrange this in such a way that equations (12) should both have at least

$$\left[\frac{1}{2} \cdot \frac{Z}{(k+1)^2 2^k}\right] = \left[\frac{Z}{(k+1)^2 2^{k+1}}\right]$$

solutions, as stated.

By our assumption every residue class mod T $(T = v-u)$ contains integers which are the sums of distinct a's. Thus by Lemma 2 there are R a's

$$a_{i_1}, \ldots, a_{i_R}, \quad R \leqslant T, \tag{13}$$

so that every residue class mod T is the sum of distinct a's from the sequence (13).

Henceforth we shall consider only those solutions of (12) where each $\underline{a}$ occurs in only one of the equations (12) and where none of the

numbers (13) occur in any of the equations (12). Under these conditions both of the equations (12) have at least

$$\left[\frac{Z}{(k+1)^2 2^{k+1}}\right]-T>\left[\frac{Z}{(k+1)^2 2^{k+1}}\right]-\frac{10n}{Z2^k}>\left[\frac{Z}{(k+1)^2 2^{k+2}}\right]=L \tag{14}$$

since by (6), (8), (10), (11) and (12) $Z>\frac{1}{4}Cn^a$ and, by Lemma 1, $2^k\leqslant\sqrt{n}$.

Clearly all the numbers

$$Lu,(L-1)u+v,\ldots,Lv;\qquad Lv=Lu+LT \tag{15}$$

can be written as the sums of distinct a's; in fact, they can be written as the distinct sums of solutions of (12) without using any of the numbers (13). The numbers (15) are $L+1$ consecutive terms of an arithmetic progression with first term Lu and difference $v-u=T$.

Now we shall show that every integer

$$Lu+sT,\qquad s\geqslant 0 \tag{16}$$

is the sum of distinct a's where the numbers (13) will not be used. If we have accomplished this, then it immediately follows from Lemma 3 that by using the integers (13) every integer not less than $Lu+\sum_{j=1}^{R}a_{ij}$ is the sum of distinct a's, and hence our Theorem is proved.

Thus we only have to prove our statement about the integers of the form (16). Denote by B the sequence $b_1<b_2<\ldots$ which we obtain by omitting from the sequence A the numbers (13) and the $\left[\frac{1}{2}\left(A(n)-A(n/2)\right)\right]$ numbers $a_{l+1},\ldots,a_{l+z}$, some (or all) of which were used in the representation of the numbers (15). Consider now the T smallest b's, $b_1,\ldots,b_T$. By lemma 3 there is a sum $b_{i_1^{(1)}}+\ldots+b_{i_r^{(1)}}=x_1\equiv 0\pmod T$, $1\leqslant i_1^{(1)}<\ldots<i_r^{(1)}\leqslant T$. Omit the b's occurring in the representation of x_1 and consider the T smallest amongst the remaining b's; again by Lemma 3 there is a sum

$$b_{i_1^{(2)}}+\ldots+b_{i_{r_2}^{(2)}}=x_2\equiv 0\pmod T,\qquad r_2\leqslant T$$

(now we can only assert $i_{r_1}^{(1)}<i_{r_2}^{(2)}\leqslant 2T$). Suppose we have already defined $x_1,\ldots,x_{k-1}$. Take the T smallest b's which do not occur in the representation of $x_1,\ldots,x_{k-1}$. By Lemma 3 the sum of some of them is a multiple of T; this defines x_k. Clearly $x_k\to\infty$ but the x_k are not necessarily monotonically increasing. Every b occurs in the representation of at most one x_k, and it clearly follows from Lemma 3 that there are fewer than T b's which never occur as summands for some x_k. Now we prove

$$x_1<LT \tag{17}$$

and for all $k>1$

$$x_1+\ldots+x_{k-1}+LT>x_k. \tag{18}$$

Assume that (17) and (18) are already proved. Then it is easy to see that every integer (16) is the sum of distinct a's where the numbers (13) are not used. First of all we have already shown that the numbers (15) (i.e. numbers of the form $Lu+sT$, $0\leqslant s\leqslant L$) are the sums of distinct a's, where the numbers (13) and the a's which occur as summands in the x_k have not been used. Thus all numbers (16) of the interval $(Lu+x_1, Lu+LT+x_1)$ are the sums of distinct a's, and by (17) this implies that all numbers (16) of the interval $(Lu, Lu+LT+x_1)$ are the sums of distinct a's. Assume that we have already shown that all numbers (16) of the interval $(Lu, Lu+LT+x_1+\ldots+x_{k-1})$ are the sums of distinct a's. Clearly all numbers (16) of $(Lu+x_k, Lu+LT+x_1+\ldots+x_{k-1}+x_k)$ are the sums of distinct a's. By (18) this implies that all numbers (16) of $(Lu, Lu+LT+x_1+\ldots+x_k)$ are the sums of distinct a's (the numbers (13) are clearly not used as summands). Thus clearly every number (16) is the sum of distinct a's, and our proof is complete.

Thus to prove our Theorem we only have to prove (17) and (18). First we show (17). By (6), (8), (10), (11), (12) and (14) we have

$$Z>\frac{D}{8}n^a,\qquad T\leqslant\frac{10n}{Z2^k}<\frac{80n^{1-a}}{D2^k},\qquad L=\left[\frac{Z}{(k+1)^2 2^{k+2}}\right]>\frac{Dn^a}{(k+1)^2 2^{k+5}} \tag{19}$$

(if $\lim A(x)/x^a=\infty$ then in (19) C should replace D).

Thus by (19) we have, for sufficiently large n,

$$A\left(\frac{n}{2}\right)>C\left(\frac{n}{2}\right)^a>2T.$$

Thus, by the definition of the b's, $b_T\leqslant a_{2T}$. Hence by the definition of x_1 and (0) we have

$$x_1\leqslant\sum_{i=1}^{T}b_i\leqslant\sum_{i=1}^{2T}a_i<\sum_{i=1}^{2T}\left(\frac{i}{C}\right)^{1/a}+c_2<2T\left(\frac{2T}{C}\right)^{1/a}+c_2.$$

Thus to show (17) we only have to show

$$2\left(\frac{2T}{C}\right)^{1/a}+c_2<L \tag{20}$$

or by (19) and (20)

$$2\left(\frac{80n^{1-a}}{D2^k}\right)^{1/a}+c_2<\frac{Dn^a}{(k+1)^2 2^{k+5}}. \tag{21}$$

But (21) clearly follows from $a^2+a=1$ for sufficiently large C and n (since the numerator of the right side is larger than the numerator of the left and the denominator is smaller and c_2 can be neglected). Thus (17) is proved.

In the proof of (18) we will often omit the simple but tedious computations (1). Denote by a_y the greatest $\underline{a}$ which occurs in the representation of $x_1, \dots, x_{k-1}$. By the definitions of x_k we have

$$x_k < T a_{y+T} . \tag{22}$$

Assume first $a_y < a_{l+1}$ ($a_{l+1}, \dots, a_{l+Z}$ were the Z largest a's in $(n/2, n)$ to which we applied lemma 1; these a's did not occur amongst the b's). Then there are at most $2T$ a's not exceeding a_y which do not occur as summands in the representation of the x_i, $1 \leqslant i \leqslant k-1$ (i.e. the $R \leqslant T$ numbers (13) and possibly T b's. Thus

$$x_1 + \dots + x_{k-1} > \sum_{i=1}^{y} a_i - 2T a_y > \sum_{i=1}^{y} a_i - 2T a_{y+T} . \tag{23}$$

Assume next that $a_y \geqslant a_{l+1}$; then we must have $a_y > a_{l+Z}$ and $a_y > n$, since the a_{l+i}, $1 \leqslant i \leqslant Z$, do not occur among the b's and thus do not occur in the representation of the x's; $a_y > n$ follows since a_{l+Z} has been the largest $\underline{a}$ not exceeding n. Here Z further a's not exceeding a_y do not occur as summands in the representation of the x_i. Thus we have

$$x_1 + \dots + x_{k-1} > \sum_{i=1}^{y-Z} a_i - 2T a_{y+T} .$$

But from (11) and $a_y > n$ we have $y > 2Z$. Thus

$$x_1 + \dots + x_{k-1} > \sum_{1 \leqslant i \leqslant (y+1)/2} a_i - 2T a_{y+T} . \tag{24}$$

Thus by (22), (23) and (24), (18) will follow if we show that

$$\sum_{1 \leqslant i \leqslant (y+1)/2} a_i + LT > 3T a_{y+T} . \tag{25}$$

(25) is trivial unless $a_{y+T} > \frac{1}{3} L$. Thus by (0) (L is large)

$$\left(\frac{y+T}{C}\right)^{1/\alpha} > \frac{L}{3} . \tag{26}$$

(26) implies, by a simple computation (as in the proof of (20)) using $\alpha^2 + \alpha = 1$ and (19), that

$$y > L^{\alpha} > T . \tag{27}$$

By (0) and (27)

$$a_{y+T} < a_{2y} < \left(\frac{2y}{C}\right)^{1/\alpha} .$$

(1) We shall constantly use the fact that L^{α}/T is large; this is implied by (19).

Thus to prove (18) it will suffice to show that if $y > L^{\alpha}$ then

$$\sum_{1 \leqslant i \leqslant (y+1)/2} a_i > 3T \left(\frac{2y}{C}\right)^{1/\alpha} . \tag{28}$$

To prove (28) we distinguish three cases. Assume first that in (5) $\beta > \alpha$. By the definition of β we infer that, for every $\varepsilon > 0$ and $t > t_0(\varepsilon)$, $a_t > t^{1/(\beta+\varepsilon)}$. Thus by a simple computation

$$\sum_{1 \leqslant i \leqslant (y+1)/2} a_i > c_3 y^{1+1/(\beta+\varepsilon)} - c_4 , \tag{29}$$

where c_3 is an absolute constant and c_4 depends only on ε. Hence we have to show that for $y > L^{\alpha}$

$$c_3 y^{1+1/(\beta+\varepsilon)} - c_4 > 3T \left(\frac{2y}{C}\right)^{1/\alpha} . \tag{30}$$

Since $1 + 1/(\beta+\varepsilon) > 1/\alpha$ for sufficiently small ε ($\beta \leqslant 1$, $\alpha = (\sqrt{5}-1)/2$), it will suffice to show (30) for $y > L^{\alpha}$, and this follows by a simple computation using (12), (14) $\alpha^2 + \alpha = 1$ and $\beta > \alpha$.

Assume next that $\beta = \alpha$ but (9) holds. Put $A(n)/n^{\alpha} = u_n$. By (10) we have

$$u_n = \frac{A(n)}{n^{\alpha}} > \frac{A(m)}{m^{\alpha}} \quad \text{for} \quad m < n . \tag{31}$$

By (31) and (11)

$$Z > \frac{u_n}{8} n^{\alpha} . \tag{32}$$

From (31) we have for $a_t \leqslant n$

$$a_t \geqslant \left(\frac{t}{u_n}\right)^{1/\alpha} \tag{33}$$

and for $a_t > n$ we infer from $\beta = \alpha$ that for every $\varepsilon > 0$ if $n > n_0(\varepsilon)$

$$a_t > t^{1/(\alpha+\varepsilon)} . \tag{34}$$

Thus if $a_{[(y+1)/2]} \leqslant n$ we have from (33)

$$\sum_{1 \leqslant i \leqslant (y+1)/2} a_i > c_5 \frac{y^{1+1/\alpha}}{u_n^{1/\alpha}} . \tag{35}$$

From (35) and (28) we only have to show that for $y > L^{\alpha}$

$$c_5 \frac{y^{1+1/\alpha}}{u_n^{1/\alpha}} > 3T \left(\frac{2y}{C}\right)^{1/\alpha} , \tag{36}$$

which again follows by a simple computation using (19) and (32) (it again suffices to show (36) for $y = L^{\alpha}$).

ACTA ARITHMETICA
VII (1962)

Finally, if $a_{[(y+1)/2]} > n$, we have by (34) for $c_6 = c_6(\varepsilon)$

$$\sum_{1 \leqslant i \leqslant (y+1)/2} a_i > c_6 y^{1+1/(\alpha+\varepsilon)}. \tag{37}$$

Thus we have to show that

$$c_6 y^{1+1/(\alpha+\varepsilon)} > 3T\left(\frac{2y}{C}\right)^{1/\alpha},$$

or, for sufficiently large C,

$$y^{1-\varepsilon/\alpha^2} > T. \tag{38}$$

Indeed (38) is trivial for sufficiently large C and n, since from

$a_{[(y+1)/2]} > n$, $y > A(n) > Cn^{\alpha}$ and by (19) we have $T < 80n^{1-\alpha}/C$.

In the third case (7) holds. By (7) we have for $t > t_0$, $a_t > \frac{1}{2}(t/D)^{1/\alpha}$. Thus

$$\sum_{1 \leqslant i \leqslant (y+1)/2} a_i > \frac{c_7}{D^{1/\alpha}} y^{1+1/\alpha} - c_8,$$

where c_7 and c_8 are absolute constants. Thus we only have to show that for $y > L^{\alpha}$

$$\frac{c_7}{D^{1/\alpha}} y^{1+1/\alpha} - c_8 > 3T\left(\frac{2y}{C}\right)^{1/\alpha}. \tag{39}$$

As before, it suffices to prove (39) for $y = L^{\alpha}$. By (8), (11) and (19), (39) follows from $\alpha^2 + \alpha = 1$ by a simple computation for sufficiently large C and n (if n is large $y = L^{\alpha}$ is also large).

Thus the proof of our Theorem is complete.

References

[1] B. J. Birch, *Note on a problem of Erdös*, Proc. Cambridge Phil. Soc. 55 (1959), pp. 370-373.

[2] J. W. S. Cassels, *On the representation of integers as the sums of distinct summands taken from a fixed set*, Acta Szeged 21 (1960), pp. 111-124.

Reçu par la Rédaction le 6. 1. 1962

Uniform distribution mod 1 (II)

by

H. Kesten (1) (Ithaca, N. Y.)

1. Introduction. Let $[a, b]$ be an interval properly contained in $[0, 1]$ and define

$$f(\xi) = \begin{cases} 1 & \text{if} \quad a \leqslant \xi \leqslant b, \\ 0 & \text{if} \quad 0 \leqslant \xi < a \quad \text{or} \quad b < \xi < 1, \end{cases} \tag{1.1}$$
$$f(\xi+1) = f(\xi).$$

$f(\xi)$ is the characteristic function of $[a, b]$ extended periodically. The present paper is concerned with the distribution of the sums

$$\sum_{k=1}^{N} f(y+kx)$$

which equal the number of terms among $y+x, y+2x, \ldots, y+Nx$ with fractional part in $[a, b]$. We assume that x and y are independent random variables each with a uniform distribution on $[0, 1]$ and show that

$$(\log N)^{-1} \sum_{k=1}^{N} \big(f(y+kx) - (b-a)\big)$$

has asymptotically a Cauchy distribution. This is expressed in the following

Theorem. *If $|E|$ denotes the Lebesgue measure of the set E, then, for every real α,*

$$\begin{aligned} &\lim_{N\to\infty} P\Big\{(\log N)^{-1} \sum_{k=1}^{N} \big(f(y+kx) - (b-a)\big) \leqslant \alpha\Big\} \\ &= \lim_{N\to\infty} \Big|\Big\{x, y \mid (\log N)^{-1} \sum_{k=1}^{N} \big(f(y+kx) - (b-a)\big) \leqslant \alpha,\ 0 \leqslant x,\ y \leqslant 1\Big\}\Big| \\ &= \frac{1}{\pi} \int_{-\infty}^{\alpha} \frac{dt}{1+t^2}. \end{aligned} \tag{1.2}$$

(1) Research supported by the National Science Foundation under project NSF-G 18837.