On a problem of Sidon in additive number theory.

By P. ERDŐS in Notre Dame (Indiana, U. S. A.)

To the memory of S. Sidon.

Let $0 < a_1 < a_2 \ldots$ be an infinite sequence of positive integers. Denote by $f(n)$ the number of solutions of $n = a_i + a_j$. About twenty years ago, SIDON[1]) raised the question wether there exists a sequence a_i satisfying $f(n) > 0$ for all $n > 1$ and $\lim f(n)/n^\varepsilon = 0$ for all $\varepsilon > 0$. In the present note, I will construct such a sequence. In fact, my sequence will satisfy

$$(1) \qquad 0 < f(n) < c_1 \log n$$

for all $n > 1$. (The c's will denote suitable positive absolute constants.)

It seems probable that (1) can not be strengthened very much. TURÁN and I conjectured[2]) that if $f(n) > 0$ for all $n > n_0$ then $\limsup f(n) = \infty$; this conjecture seems very hard to prove. A still stronger conjecture would be the following: Let $a_1 < a_2 < \cdots$ be an infinite sequence of integers satisfying $a_k < ck^2$. Then $\limsup f(n) = \infty$.

We will not be able to construct our sequences explicitely, but we will be able to show that in some sense almost all sequences satisfy (1). The idea of our construction is the following. Define A_k as the least integer greater than $c_2 k^{1/2} 2^{k/2}$. Pick in all possible ways A_k integers from the interval $(2^k, 2^{k+1})$. One can do this in $\binom{2^k}{A_k}$ ways. One thus obtains the integers $b_1^{(k)} < b_2^{(k)} < \cdots < b_{A_k}^{(k)}$. I will show that if for each k one neglects $o\binom{2^k}{A_k}$ "bad" choices of the $b_i^{(k)}$ and forms a sequence $a_1 < a_2 < \cdots$ from any of the "good" choices of the $b_i^{(k)}$ $(k = 1, 2, \ldots;\ 1 \leqq i \leqq A_k)$ the sequence $a_1 < a_2 < \cdots$ will satisfy (1). Thus roughly speaking (1) will be satisfied for almost all choices of the $b_i^{(k)}$. We have to make one more remark. For small values of k it may happen that $c_2 2^{k/2} k^{1/2} > 2^k$. In this case the $b_i^{(k)}$ are simply all the integers of the interval $(2^k, 2^{k+1})$. Also, it is clearly enough to show that $f(n) > 0$ for all

[1]) Oral communication.

[2]) P. ERDŐS and P. TURÁN, On a problem of Sidon in additive number theory, and on some related problems, *Journal London Math. Soc.*, **16** (1941), 212—215.

$n > n_0$, for if $f(n) > 0$ for all $n > n_0$ we can simply adjoin all $n < n_0$ to the a's and the new sequence satisfies $f(n) > 0$ for all $n > 1$ and $f(n) < c_1 \log n + n_0 < c_3 \log n$.

First we state and prove some simple inequalities on binomial coefficients which we will need later.

Let u, v, x and l be integers, $x < u$ and $x < v$. Denote by l_1 the greatest integer less than $ux/8(u+v)$ and by $2l_2$ the least even integer greater than $4rux/(u+v)$ $(r>1)$. We have

$$\sum_{l=0}^{l_1} \binom{u}{l}\binom{v}{x-l} < \frac{2}{2^{l_1}}\binom{u+v}{x}, \tag{2}$$

$$\sum_{l=2l_2}^{x} \binom{u}{l}\binom{v}{x-l} < \frac{1}{1-\frac{1}{r}} \cdot \frac{1}{r^{l_2}}\binom{u+v}{x}. \tag{3}$$

A simple computation shows that for $l \leqq 2l_1$

$$\binom{u}{l}\binom{v}{x-l} < \frac{1}{2}\binom{u}{l+1}\binom{v}{x-l-1}.$$

Thus

$$\sum_{l=0}^{l_1} \binom{u}{l}\binom{v}{x-l} < 2\binom{u}{l_1}\binom{v}{x-l_1} < \frac{2}{2^{l_1}}\binom{u}{2l_1}\binom{v}{x-2l_1} <$$

$$< \frac{2}{2^{l_1}}\sum_{l=0}^{x}\binom{u}{l}\binom{v}{x-l} = \frac{2}{2^{l_1}}\binom{u+v}{x},$$

which proves (2). Again by a simple computation we have for $l \geqq l_2$

$$\binom{u}{l}\binom{v}{x-l} > r\binom{u}{l+1}\binom{v}{x-l-1}.$$

Thus

$$\sum_{l=2l_2}^{x} \binom{u}{l}\binom{v}{x-l} < \frac{1}{1-\frac{1}{r}}\binom{u}{2l_2}\binom{v}{x-2l_2} < \frac{1}{\left(1-\frac{1}{r}\right)r^{l_2}}\binom{u}{l_2}\binom{v}{x-l_2} <$$

$$< \frac{1}{\left(1-\frac{1}{r}\right)r^{l_2}}\sum_{l=0}^{x}\binom{u}{l}\binom{v}{x-l} < \frac{1}{\left(1-\frac{1}{r}\right)r^{l_2}}\binom{u+v}{x},$$

which proves (3).

Now we can start to prove our Theorem. First of all, we show that for all but $o\binom{2^k}{A_k}$ choices of the $b_i^{(k)}$ $(k = 1, 2, \ldots; 1 \leqq i \leqq A_k)$, $f(n) > 0$ for all $n > 1$. Assume that we already have chosen the $b_i^{(l)}$ $(1 \leqq l < k;\ 1 \leqq i \leqq A_l)$ so that $f(n) > 0$ for all $n < 2^{k-1} + 2^k$. Then we show that for all but $o\binom{2^k}{A_k}$ choices of the $b_i^{(k)}$ $(1 \leqq i \leqq A_k)$, we have $f(n) > 0$ for all n satisfying $2^{k-1} + 2^k \leqq n < 2^k + 2^{k+1}$.

Assume first

$$2^{k-1}+2^k \leqq n < 2^{k-2}+2^{k+1}. \tag{4}$$

If $f(n)=0$ then clearly none of the integers

$$n-b_j^{(k-2)}\ (1 \leqq j \leqq A_{k-2})$$

could be $b_i^{(k)}$'s. Since by construction $2^{k-2} \leqq b_j^{(k-2)} < 2^{k-1}$, all the $n-b_j^{(k-2)}$ are in $(2^k, 2^{k+1})$. Thus the number of possible choices of $b_i^{(k)}$ $(1 \leqq i \leqq A_k)$ for which $f(n)=0$ is certainly not greater than

$$\binom{2^k-A_{k-2}}{A_k}.$$

Thus the number of possible choices of the $b_i^{(k)}$ $(1 \leqq i \leqq A_k)$ so that $f(n)=0$ for at least one n satisfying (4) is less than

$$2^k\binom{2^k-A_{k-2}}{A_k} = 2^k\binom{2^k}{A_k}\frac{(2^k-A_k)(2^k-A_k-1)\cdots(2^k-A_k-A_{k-2}+1)}{2^k(2^k-1)\cdots(2^k-A_{k-2}-1)} <$$

$$< 2^k\binom{2^k}{A_k}\left(1-\frac{A_k}{2^k}\right)^{A_{k-2}} < 2^k\binom{2^k}{A_k}\exp\left(-\frac{A_kA_{k-2}}{2^k}\right) < 2^k\binom{2^k}{A_k}\exp\left(-\frac{c_2^2}{4}k\right) = o\binom{2^k}{A_k}$$

for $c_2>4$, which proves our assertion if (4) holds.

Assume next

$$2^{k-2}+2^{k+1} \leqq n < 2^k+2^{k+1}. \tag{5}$$

If $2^k \leqq x < 2^{k-3}+2^k$ then only one of the numbers x and $n-x$ can be a $b_i^{(k)}$ (otherwise $f(n)>0$). Assume that there are l $b_i^{(k)}$'s in $(2^k, 2^{k-3}+2^k)$ where l can take the values $0, 1, \ldots, A_k$. Then by what has been just said none of the numbers $n-b_i^{(k)}$ $(1 \leqq i \leqq l)$ can be $b_i^{(k)}$'s [these numbers $n-b_i^{(k)}$ $(1 \leqq i \leqq l)$ are all in $(2^{k-3}+2^k, 2^{k+1})$]. Thus the number of possible choices of the $b_i^{(k)}$ $(1 \leqq i \leqq A_k)$ so that $f(n)=0$ for a fixed n satisfying (5) is less than or equal to

$$\sum_{l=0}^{A_k}\binom{2^{k-3}}{l}\binom{2^k-2^{k-3}-l}{A_k-l} = \Sigma_1+\Sigma_2 \tag{6}$$

where in Σ_1 $l<A_k/64$ and in Σ_2 $l \geqq A_k/64$.

We obtain from (2) by putting $u=2^{k-3}$, $v=2^k-2^{k-3}$, $x=A_k$, $l_1=[A_k/64]$:

$$\sum\nolimits_1 < \sum_{l<A_k/64}\binom{2^{k-3}}{l}\binom{2^k-2^{k-3}}{A_k-l} < \frac{2}{2^{A_k/64}}\binom{2^k}{A_k} = o\left(\frac{1}{2^k}\right)\binom{2^k}{A_k}. \tag{7}$$

Further

$$\sum\nolimits_2 = \sum_{l \geqq A_k/64}\binom{2^{k-3}}{l}\binom{2^k-2^{k-3}-l}{A_k-l} < \sum_{l=0}^{A_k}\binom{2^{k-3}}{l}\binom{2^k-2^{k-3}}{A_k-l}\left(1-\frac{A_k}{64\cdot 2^k}\right)^{A_k/64}.$$

The last step follows from $\binom{u}{t}\left(1-\frac{l}{u}\right)^l > \binom{u-l}{t}$ (in our case $u=2^k-2^{k-3}$, $l \geqq A_k/64$). Thus for $c_2>8$

$$\sum\nolimits_2 < \binom{2^k}{A_k}\left(1-\frac{A_k}{64\cdot 2^k}\right)^{A_k} < \binom{2^k}{A_k}\exp\left(-\frac{A_k^2}{64\cdot 2^k}\right) = o\left(\frac{1}{2^k}\right)\binom{2^k}{A_k}. \tag{8}$$

From (7) and (8) we have $\Sigma_1 + \Sigma_2 = o\left(\frac{1}{2^k}\right)\binom{2^k}{A_k}$. Thus, if (5) holds, then for all but $o\binom{2^k}{A_k}$ choices of the $b_i^{(k)}$ $(1 \leqq i \leqq A_k)$ we have $f(n) > 0$.

Since (4) and (5) cover the interval $(2^{k-1} + 2^k, 2^k + 2^{k+1})$, the proof of our assertion is complete, that is if we form a sequence $a_1 = 1 < a_2 < \cdots$ from the "good" choices of the $b_i^{(k)}$ $(k = 0, 1, \ldots; \; 1 \leqq i \leqq A_k)$, then we have $f(n) > 0$ for all $n > 1$.

Now to complete our proof we have to show that for all but $o\binom{2^k}{A_k}$ choices of the $b_i^{(k)}$ $(1 \leqq i \leqq A_k, 1 \leqq k < \infty)$ we have

$$f(n) < c_1 \log n. \tag{9}$$

To show (9) it will clearly suffice to show that except for $o\binom{2^k}{A_k}$ choices of the $b_i^{(k)}$ for all n the number of solutions of

$$n = b_i^{(k)} + b_j^{(r)} \; (1 \leqq i \leqq A_k, \; 1 \leqq j \leqq A_r, \; 1 \leqq r \leqq k-1) \tag{10}$$

is less than $c_1 \log n/2$, and that except for $o\binom{2^k}{A_k}$ choices of the $b_i^{(k)}$ the number of solutions of

$$n = b_{i_1}^{(k)} + b_{i_2}^{(k)} \; (1 \leqq i_1, i_2 \leqq A_k) \tag{11}$$

is for all n less than $c_1 \log n/2$. (A simple argument shows that (10) and (11) imply (9)).

First we deal with (10). We clearly can assume that

$$2^k < n \leqq 2^k + 2^{k+1}$$

(i. e. $b_j^{(r)} \leqq 2^k$, $2^k < b_i^{(k)} \leqq 2^{k+1}$). Consider the numbers $n - b_j^{(r)}$ $(1 < j < A_r$, $1 \leqq r \leqq k-1)$. Suppose that z of them are in the interval $(2^k, 2^{k+1})$. We evidently have

$$z \leqq A_1 + A_2 + \cdots + A_{k-1} < 4c_2 k^{1/2} 2^{k/2}.$$

If (10) is to have not less than $c_1 \log n/2$ solutions, then at least $c_1 \log n/2$ of these numbers $n - b_j^{(r)}$ must be $b_i^{(k)}$'s. Thus since $c_1 \log n/2 > c_1 k/4$, the number of possible choices of the $b_i^{(k)}$ for which (10) has more than $c_1 \log n/2$ solutions is not greater than

$$\sum_{l > c_1 k/4} \binom{z}{l}\binom{2^k - z}{A_k - l}.$$

Now use (3) with $x = A_k, u = z, v = 2^k - z, l_2 = c_1 k/8$. Further by the definition of r

$$\left[\frac{4rzA_k}{2^k}\right] = \left[\frac{c_1 k}{4}\right] \quad \text{or} \quad r > 2 \quad \text{for} \quad c_1 > 126 c_2^2.$$

Thus by (3) for $c_1 > 8$

$$\sum_{l > c_1 k/4} \binom{z}{l}\binom{2^k - z}{A_k - l} < \frac{2}{2^{c_1 k/8}}\binom{2^k}{A_k} = o\left(\frac{1}{2^k}\right)\binom{2^k}{A_k}.$$

Thus the number of choices of the $b_i^{(k)}$ for which (10) has for at least one n $(2^k < n < 2^k + 2^{k+1})$ more than $c_1 \log n/2$ solutions is $o\binom{2^k}{A_k}$ as stated.

To complete our proof we finally investigate (11). Here we clearly can assume $2^{k+1} \leqq n < 2^{k+2}$. If for some n (11) has more than $c_1 \log n/2$ solutions we must have

$$n = b_{i_1}^{(k)} + b_{i_2}^{(k)} = b_{i_3}^{(k)} + b_{i_4}^{(k)} = \cdots = b_{i_{2t-1}}^{(k)} + b_{i_{2t}}^{(k)}, \qquad t = \left[\frac{c_1 k}{4}\right] (< c_1 \log n/2).$$

(We can ignore the solution $2b_i^{(k)} = n$.) There clearly are $\binom{2^k}{t}$ possible choices for the $b_{2r-1}^{(t)}$, $1 \leqq r \leqq t$; the $b_{2r}^{(t)}$ are then determined, and for the remaining $b_i^{(k)}$ there are $\binom{2^k - 2t}{A_k - 2t}$ possible choices. Thus, if the number of possible choices for the $b_i^{(k)}$ $(1 \leqq i \leqq A_k)$ for which (11) has, for at least one n, $2^{k+1} \leqq n < 2^{k+2}$, more than $c_1 \log n/2$ solutions, is not greater than

$$2^{k+1}\binom{2^k}{t}\binom{2^k - 2t}{A_k - 2t} < 2^{k+1}\frac{2^{kt}}{t!}\left(\frac{2^k}{A_k}\right)^{-2t}\binom{2^k}{A_k} =$$
$$= \binom{2^k}{A_k}\frac{2^{k+1}}{t!}\,\frac{A_k^{2t}}{2^{kt}} < \binom{2^k}{A_k}\frac{2^{k+1}}{t!}(c_2^2 k)^k = o\binom{2^k}{A_k}$$

for sufficiently large c_1, $t = \left[\frac{c_1 k}{4}\right]$. Thus our theorem is completely proved.

UNIVERSITY OF NOTRE DAME.

(Received February 22, 1954.)