JOURNAL OF NUMBER THEORY 18, 99–109 (1984)

On Disjoint Sets of Differences

P. ERDÖS

Mathematical Institute of the Academy, Budapest, Realtanoda u. 13–15, H-1053, Budapest, Hungary

AND

ROBERT FREUD

University Eötvös Loránd, Department of Algebra and Number Theory, Budapest, Muzeum krt. 6–8., H-1088, Budapest, Hungary

Communicated by H. Zassenhaus

Received January 20, 1982

We investigate integer sequences A and B where $\{A-A\} \cap \{B-B\} = 0$. We solve a problem of P. Erdös and R. L. Graham and prove several results on the behaviour of $A(x)B(x)/x$, $A(x)/\sqrt{x}$ and $B(x)/\sqrt{x}$.

Sidon's problems are of central interest in combinatorial number theory (see, e.g., [1; 2, pp. 48–49; 3, Chap. II]). An infinite sequence A of positive integers is called a Sidon sequence, if the differences $a_i - a_j$ $(i \neq j)$ are all distinct. It was proved by Erdös that for a Sidon sequence

$$\liminf_{x\to\infty} \frac{A(x)}{\sqrt{x}} = 0, \qquad \text{moreover} \qquad \liminf_{x\to\infty} \frac{A(x)}{\sqrt{x/\log x}} < \infty \tag{i}$$

must hold, where $A(x)$ denotes the number of elements of A up to x.

It is quite natural to ask how much the situation changes if we cut A into two parts, A' and A'', and demand only that no $a_i' - a_j'$ should coincide with any $a_i'' - a_j''$. This question was proposed by Erdös and Graham in [2], and it seemed likely that no considerable increase can be achieved in the density of A. We shall show, however, that the situation changes dramatically, and we can construct very dense sequences.

Let us see first the precise formulation of the problem [2, p. 50]: "Let

0022-314X/84 $3.00

$A = \{a_1 < a_2 < \cdots\}$ and $B = \{b_1 < b_2 < \cdots\}$ be sequences of integers satisfying $A(x) > \varepsilon x^{1/2}$, $B(x) > \varepsilon x^{1/2}$ for some $\varepsilon > 0$. Is it true that

$$a_i - a_j = b_k - b_l \tag{1}$$

has infinitely many solutions?"

The negative answer is provided, e.g., by the following A and B; we write the numbers in binary scale, and select for A those which contain only even powers of two, and for B those which contain only odd powers of two,

$$A = \left\{ \sum_{l=0}^{n} c_{2l} 2^{2l},\, c_{2l} = 0 \text{ or } 1,\, n = 0, 1, 2, \ldots \right\},$$

$$B = \left\{ \sum_{l=0}^{n} c_{2l+1} 2^{2l+1},\, c_{2l+1} = 0 \text{ or } 1,\, n = 0, 1, 2, \ldots \right\}.$$

Then (1) is possible only in the trivial case, since it is equivalent to

$$a_i + b_l = a_j + b_k \tag{2}$$

and every integer can be uniquely written as the sum of different powers of two. On the other hand

$$\liminf_{x \to \infty} \frac{\min\{A(x), B(x)\}}{\sqrt{x}} = 1/\sqrt{2}$$

(cf. (i)!), since the "worst" case occurs just before a new digit turns up in B;

$$B(2^{2s-1} - 1) = 2^{s-1} \sim \frac{1}{\sqrt{2}} \cdot \sqrt{2^{2s-1} - 1}.$$

This settles the original question in the negative (for $\varepsilon = 1/\sqrt{2}$).

In the following we consider such sequences A and B where (1) (or (2)) has only trivial solutions, and investigate the behaviour of $A(x)\,B(x)/x$, $A(x)/\sqrt{x}$ and $B(x)/\sqrt{x}$.

We introduce some notations:

$$SP = \limsup_{x \to \infty} \frac{X(x)\,B(x)}{x},$$

$$IP = \liminf_{x \to \infty} \frac{A(x)\,B(x)}{x},$$

$$SN = \limsup_{x \to \infty} \frac{\min\{A(x), B(x)\}}{\sqrt{x}},$$

$$IN = \liminf_{x\to\infty} \frac{\min\{A(x), B(x)\}}{\sqrt{x}},$$

$$SX = \limsup_{x\to\infty} \frac{\max\{A(x), B(x)\}}{\sqrt{x}},$$

$$IX = \liminf_{x\to\infty} \frac{\max\{A(x), B(x)\}}{\sqrt{x}}$$

(S stands for lim sup, I for lim inf, P for product, N for min and X for max).

It is easy to check that in our previous example

$$SP = 3/2, \qquad IP = 1,$$
$$SN = \sqrt{3}/\sqrt{2}, \qquad IN = 1/\sqrt{2},$$
$$SX = \sqrt{3}, \qquad IX = 1.$$

THEOREM 1. *The largest possible value of SP is 2, moreover the following more precise estimations hold*:

1.1. *To any function $H(x)$ with $\limsup_{x\to\infty} H(x) = \infty$, we can construct A and B so that*

$$A(x)\,B(x) \geqslant 2x - H(x) \tag{3}$$

is valid for infinitely many (integer) values of x.

1.2. *The previous result is best possible: for any A and B, $A(x)\,B(x) - 2x \to -\infty$ $(x \to \infty)$.*

THEOREM 2.

2.1. *$\frac{5}{2}IP + 2SP \leqslant 7$, in particular $IP \leqslant 14/9$.*

2.2. *$IP + \frac{3}{2}SP \leqslant 4$, in particular $SP = 2$ implies $IP \leqslant 1$.*

Remark. We could not yet decide if $IP > 1$ is possible at all.

THEOREM 3.

3.1. *The largest possible value of SN is $\sqrt{2}$, that of IX is ∞.*

3.2. *$IN > 1/\sqrt[4]{2} - \varepsilon$ is attainable for any $\varepsilon > 0$.*

3.3. *To any $\varepsilon > 0$ we can construct an A and B with $SP > 2 - \varepsilon$ and $IN > 0$, $SX < \infty$ but $SP = 2$ implies $IN = 0$ and $SX = \infty$.*

Remark. 2.1 and 3.2 imply that the largest possible value of IN lies between $1/\sqrt[4]{2}$ and $\sqrt{14/9}$, but we have no better estimations yet.

THEOREM 4. *If $IN > 0$, then neither $A(x)/\sqrt{x}$ nor $B(x)/\sqrt{x}$ can tend to a limit.*

We shall consider further generalizations in a next paper.

Proofs. We shall frequently use the following generalization of the example in the Introduction. We write the numbers by the help of a generalized number system, and put into A those numbers where the even digits are zero, and into B those ones where the odd digits are zero. Formally: let $k_1, k_2, \ldots, k_m, \ldots$ be arbitrary integers greater than one, and

$$A = \{c_0 + c_2 k_1 k_2 + \cdots + c_{2s} k_1 k_2 \cdots k_{2s}, \quad 0 \leqslant c_{2i} \leqslant k_{2i+1} - 1, s = 0, 1, 2, \ldots\},$$

$$(*)$$

$$B = \{c_1 k_1 + c_3 k_1 k_2 k_3 + \cdots + c_{2s-1} k_1 k_2 \cdots k_{2s-1}, \quad 0 \leqslant c_{2i-1} \leqslant k_{2i} - 1, s = 1, 2, \ldots\}.$$

Clearly (2) is possible only in the trivial case.

We mention that for any A and B of this type we have $IP = 1$, since there are exactly $A(x) B(x)$ numbers of the form $a_i + b_t$ with $a_i \leqslant x$ and $b_t \leqslant x$, and so before a new digit turns up in A or in B, $A(x) B(x) = x + 1$ (for $x = k_1 k_2 \cdots k_j - 1$).

The original example is the special case $k_1 = k_2 = \cdots = 2$.

Proof of Theorem 1. We may assume $a_1 = b_1 = 0$, and then $a_i \neq b_j$ for $i, j > 1$.

$A(x) B(x) \leqslant 2x$ is obvious, since for $a_i \leqslant x$, $b_t \leqslant x$, $0 \leqslant a_i + b_t \leqslant 2x - 1$, and all the numbers $a_i + b_t$ are distinct.

To prove 1.2, we assume indirectly that for some c, $A(x) B(x) \geqslant 2x - c$ infinitely often. For any such x, there exists a sum $a_i + b_t \geqslant 2x - c$, where $a_i \leqslant x$, $b_t \leqslant x$. Then $a_i \geqslant x - c$ and $b_t \geqslant x - c$ must hold as well, and so

$$|a_i - b_t| \leqslant c. \tag{4}$$

But (2) is clearly equivalent to

$$a_i - b_k = a_j - b_t, \tag{5}$$

i.e., all the differences $a_i - b_k$ are distinct, and so (4) cannot be valid infinitely often, which is a contradiction.

To show 1.1 we take the construction $(*)$, and calculate $A(x) B(x)$ for

$$x = k_1 k_2 \cdots k_{2s} + (k_{2s-1} - 1) k_1 k_2 \cdots k_{2s-2} + (k_{2s-3} - 1) k_1 k_2 \cdots k_{2s-4} + \cdots + (k_1 - 1).$$

Now all those numbers can be written in the form $a_i + b_t$ with $a_i \leqslant x$, $b_t \leqslant x$, which have $2s+1$ digits and their first digit is 0 or 1. Hence $A(x)\,B(x)=2k_1k_2\cdots k_{2s}$.

On the other hand $x\leqslant k_1k_2\cdots k_{2s}+k_1k_2\cdots k_{2s-1}$. Thus if k_{2s} is large enough then $A(x)\,B(x)$ is "nearly" $2x$, and (3) can be easily guaranteed.

We mention that we can prove 1.1 also by an alternative version of construction (*), which is an iterative process. We sketch it briefly as follows. Assume that we have already constructed A and B till x_n, the largest value of A and B is x_n and x_n-y_n, respectively, and all numbers up to $2x_n-y_n$ can be uniquely expressed as a_i+b_j, i.e., $A(x_n)\,B(x_n)=2x_n-y_n+1=v$. Now we translate A by $v, 2v, \ldots, (r_n-1)v$ and B by r_nv. Then the largest value of B is x_{n+1}, that of A is $x_{n+1}-y_{n+1}$, where

$$x_{n+1}=r_n(2x_n-y_n+1)+(x_n-y_n)$$

and

$$y_{n+1}=2x_n-2y_n+1,$$

and all numbers up to $2x_{n+1}-y_{n+1}$ can be uniquely written in the form a_i+b_j. Since y_{n+1} does not depend on r_n, we can easily guarantee (3).

Proof of Theorem 3. 3.1. $SP\leqslant 2$ shows that $SN\leqslant\sqrt{2}$. To prove the possibility of equality we consider the (*) construction used in the proof of Theorem 1. For the x there,

$$A(x)=2k_{2s-1}k_{2s-3}\cdots k_1$$

and

$$B(x)=k_{2s}k_{2s-2}\cdots k_2$$

(the ith digit from the right can take k_i values with the exception of the $2s+1$st digit, which can be just 0 or 1).

With the suitable choice of the k_i's we can clearly assure both $A(x)=B(x)$ and the "very big" value of k_{2s} (the latter is necessary for $A(x)\,B(x)\sim 2x$).

To make IX large, we choose the k_{2i-1} values to be greater than the k_{2i} values, and so $A(x)$ will "dominate" $B(x)$.

We can also determine the extremal order of magnitude of $A(x)$. The previous argument shows the possibility of $A(x)/x$ tending to 0 arbitrarily slowly. On the other hand it is obvious that $\lim_{x\to\infty}A(x)/x=0$, if B is infinite: using $A(x)\,B(x)\leqslant 2x$ we obtain

$$\frac{A(x)}{x}\leqslant\frac{2}{B(x)}.$$

3.2. Let p/q be a rational number, $1/\sqrt{2}-\varepsilon < p/q < 1/\sqrt{2}$. Put $k_1 = p$, $k_2 = q$, $k_3 = k_4 = \cdots = 2$. Then for

$$x = k_1 k_2 \cdots k_{2s} - 1 = pq \cdot 2^{2s-2} - 1,$$
$$A(x) = k_1 k_3 \cdots k_{2s-1} = p \cdot 2^{s-1},$$
$$B(x) = k_2 k_4 \cdots k_{2s} = q \cdot 2^{s-1},$$

thus

$$\frac{\min\{A(x), B(x)\}}{\sqrt{x}} \sim \sqrt{\frac{p}{q}} > \frac{1}{\sqrt[4]{2}} - \varepsilon.$$

Similarly, for

$$x = k_1 k_2 \cdots k_{2s+1} - 1 = 2p \cdot q \cdot 2^{2s-2} - 1,$$
$$A(x) = k_1 k_3 \cdots k_{2s+1} = 2p \cdot 2^{s-1},$$
$$B(x) = k_2 k_4 \cdots k_{2s} = q \cdot 2^{s-1},$$

so

$$\frac{\min\{A(x), B(x)\}}{\sqrt{x}} \sim \sqrt{\frac{q}{2p}} > \frac{1}{\sqrt[4]{2}}.$$

Since these values of x are the "worst" ones from the point of view of IN, we obtain the statement.

We can easily check that this is the best possible value for IN using the $(*)$ construction. We know that for $x = k_1 k_2 \cdots k_s - 1$, $A(x)\,B(x) = x + 1$. Further, between $k_1 \cdots k_s$ and $k_1 k_2 \cdots k_{s+1} \geqslant 2k_1 k_2 \cdots k_s$ either A or B has no elements, say, A. Then denoting IN by c, we have on the one hand

$$A(x) = A(2x) \geqslant (c-\varepsilon)\sqrt{2x},$$

and on the other hand

$$A(x) \leqslant \frac{x}{B(x)} \leqslant \frac{1}{c-\varepsilon}\sqrt{x},$$

i.e.,

$$\frac{1}{c} \geqslant c\sqrt{2}$$

or

$$c \leqslant \frac{1}{\sqrt[4]{2}}.$$

3.3. Put $k_1 = k_2 = k_3 = \cdots = k$ with a big k. Then similarly to the previous calculations

$$SP = \frac{2(k+1)}{k+2}, \qquad IN = \frac{1}{\sqrt{k}} \qquad \text{and obviously } SX \cdot IN \leqslant SP,$$

i.e., $SP > 2 - \varepsilon$, $IN > 0$, and $SX < \infty$.

Assume now $SP = 2$. First we prove $IN = 0$. Assume indirectly, that for some positive c,

$$A(x) > c\sqrt{x} \qquad \text{and} \qquad B(x) > c\sqrt{x} \tag{6}$$

always hold. Then also

$$B(x) \leqslant 2x/A(x) < \frac{2}{c}\sqrt{x} \qquad \text{and} \qquad A(x) \leqslant 2x/B(x) < \frac{2}{c}\sqrt{x} \tag{7}$$

are valid. Let ε be very small. We take an x, for which

$$A(2x)\,B(2x) > (4-\varepsilon)x$$

is true. This means that with the exception of at most εx numbers all numbers in $[0, 4x]$ can be written in the form $a_i + b_i$, with $a_i \leqslant 2x$ and $b_i \leqslant 2x$. Clearly we can use only $a_i \leqslant x$ and $b_i \leqslant x$ for the numbers in $[0, x]$ and only $a_i > x$ and $b_i > x$ for those in $(3x, 4x]$.

Denote the elements of A and B in $[0, x]$ and in $(x, 2x]$ by A_1, B_1, A_2 and B_2, respectively. Hence

$$A_1 B_1 + A_2 B_2 > (2-\varepsilon)x \tag{8}$$

and also

$$A_2 B_2 > (1-\varepsilon)x, \qquad A_1 B_1 > (1-\varepsilon)x. \tag{9}$$

On the other hand consider now differences $a_i - b_i$. Since these must all be distinct, there are at most $2x$ of them with

$$|a_i - b_i| \leqslant x. \tag{10}$$

If a_i and b_i are both in $[0, x]$ or both in $(x, 2x]$, then (10) holds, thus

$$A_1 B_1 + A_2 B_2 \leqslant 2x. \tag{11}$$

Moreover, using (8), we obtain that there are at most εx other pairs of $a - s$ and $b - s$ which satisfy (10).

Put $d = c^4/16$. Denote by A', B', A^* and B^* the elements of A and B in $[dx, x]$ and $(x, (1+d)x]$, respectively. We show that

$$A'B^* + A^*B' > \varepsilon x, \tag{12}$$

which is a contradiction, since this means a too large number of further differences satisfying (10).

Using (7) for dx we obtain

$$A(dx) < \frac{2}{c}\sqrt{dx} = \frac{c}{2}\sqrt{x}$$

and similarly

$$B(dx) < \frac{c}{2}\sqrt{x}.$$

Combining this with (6) we have

$$A' > \frac{c}{2}\sqrt{x} \qquad \text{and} \qquad B' > \frac{c}{2}\sqrt{x}. \tag{13}$$

On the other hand

$$A\{(1+d)x\}\, B\{(1+d)x\} > (1+d-\varepsilon)x, \tag{14}$$

since we know that nearly all numbers also in $[0, (1+d)x]$ can be written in the form $a_i + b_i$, and here obviously $a_i \leqslant (1+d)x$ and $b_i \leqslant (1+d)x$. Further, combining (9) and (11) we obtain

$$A_1 B_1 < (1+\varepsilon)x. \tag{15}$$

Using (14) and (15) we infer

$$(A_1 + A^*)(B_1 + B^*) > (1+d-\varepsilon)x = (1+\varepsilon)x + (d-2\varepsilon)x$$
$$> A_1 B_1 + (d-2\varepsilon)x.$$

Hence

$$A^*B_1 + A_1B^* + A^*B^* > (d-2\varepsilon)x. \tag{16}$$

We show that

$$\max(A^*, B^*) > \left(1 - \frac{dc^2}{16}\right) \cdot \frac{dc}{4} \cdot \sqrt{x} = \frac{dcu}{4} \cdot \sqrt{x}. \tag{17}$$

If this were not true, then

$$A^*B^* < \frac{c^2d^2u^2}{16} \cdot x,$$

$$A_1B^* + A^*B_1 < 2 \cdot \frac{2}{c} \cdot \frac{dcu}{4} \cdot x = \left(1 - \frac{dc^2}{16}\right) dx,$$

i.e., $A^*B^* + A_1B^* + A^*B_1 < dx(1-u')$, which is a contradiction to (16) for ε small enough.

Finally, (17) and (13) imply (12) and this completes the proof of $IN = 0$.

To show $SX = \infty$ we can use the previous proof. We saw that if $A(2x)\,B(2x) > (4-\varepsilon)x$, then

$$A(x)\,B(x) > (1-\varepsilon)x, \tag{18}$$

and not all of the following four inequalities can hold simultaneously, for a fixed positive c, $d = c^4/16$ and for ε small enough:

$$A(x) > c\sqrt{x},$$

$$B(x) > c\sqrt{x},$$

$$A(dx) < \frac{2}{c}\sqrt{dx},$$

$$B(dx) < \frac{2}{c}\sqrt{dx}.$$

If, e.g., the third inequality is violated, this means directly that $A(dx)/\sqrt{dx}$ is large.

If, e.g., the first inequality is false, then (18) implies that $B(x) > ((1-\varepsilon)/c)\sqrt{x}$, i.e., $B(x)/\sqrt{x}$ is large.

Thus in any case $SX = \infty$.

Proof of Theorem 2. 2.1. We take an x for which

$$A(4x)\,B(4x) \geqslant 4x(SP-\varepsilon). \tag{19}$$

By assumption

$$A(2x)\,B(2x) \geqslant 2x(IP-\varepsilon) \tag{20}$$

and

$$A(3x)\,B(3x) \geqslant 3x(IP-\varepsilon). \tag{21}$$

We denote the number of elments of A and B in the intervals $((i-1)x, ix]$ by A_i and B_i, respectively, $i=1, 2, 3, 4$.

Consider the sums a_i+b_j, where $a_i \leqslant 3x$ and $b_j \leqslant 3x$. The number of these sums is $A(3x)\,B(3x)$, and at least $A(3x)\,B(3x)-4x$ of them are greater than $4x$, and for these ones both a_i and b_j are greater than x, and not both are less than $2x$. This means that

$$A_2B_3+A_3B_2+A_3B_3 \geqslant A(3x)\,B(3x)-4x \geqslant 3x(IP-\varepsilon)-4x. \quad (22)$$

Repeating the argument for $a_i+b_j>6x$, where $a_i \leqslant 4x$, $b_j \leqslant 4x$, we obtain

$$A_3B_4+A_4B_3+A_4B_4 \geqslant A(4x)\,B(4x)-6x \geqslant 4x(SP-\varepsilon)-6x. \quad (23)$$

On the other hand there are at most $4x$ differences a_i-b_j where

$$|a_i-b_j| \leqslant 2x,$$

i.e., the sum of the left-hand sides of (20), (22) and (23) is at most $4x$. So taking the sum of (20), (22) and (23) we obtain

$$4x \geqslant 2x(IP-\varepsilon)+3x(IP-\varepsilon)-4x+4x(SP-\varepsilon)-6x,$$

and since ε can be arbitrarily small, this completes the proof.

2.2. We now take an x for which

$$A(3x)\,B(3x) \geqslant 3x(SP-\varepsilon) \quad (24)$$

and using (20) and (24) we argue similarly as before.

Proof of Theorem 4. Assume indirectly that $\lim_{x\to\infty} A(x)/\sqrt{x}=c_1>0$, and $\liminf_{x\to\infty} B(x)/\sqrt{x}=c_2>0$.

Take a large but fixed k, and a very large x. We denote the number of elements of A and B in the intervals $(i-1)x, ix]$ by A_i and B_i, respectively, $i=1, 2,\ldots, k$, and put $S_i=B(ix)=B_1+B_2+\cdots+B_i$.

Since there are at most $2x$ differences where $|a_i-b_j| \leqslant x$, therefore

$$\sum_{i=1}^{k} A_iB_i \leqslant 2x.$$

On the other hand we shall show that this is false.

If x is large enough, then

$$A_i=A(ix)-A\{(i-1)x\} \sim c_1\sqrt{ix}-c_1\sqrt{(i-1)x} \sim c_1\sqrt{x}/2\sqrt{i}.$$

Hence

$$\sum_{i=1}^{k} A_i B_i \sim \frac{c_1\sqrt{x}}{2}\sum_{i=1}^{k}\frac{B_i}{\sqrt{i}} = \frac{c_1\sqrt{x}}{2}\sum_{i=1}^{k}\frac{S_i - S_{i-1}}{\sqrt{i}}$$

$$\sim \frac{c_1\sqrt{x}}{2}\sum_{i=1}^{k} S_i\left\{\frac{1}{\sqrt{i}} - \frac{1}{\sqrt{i+1}}\right\} \sim \frac{c_1\sqrt{x}}{4}\sum_{i=1}^{k}\frac{S_i}{i^{3/2}}$$

$$\geqslant \sim \frac{c_1\sqrt{x}}{4}\sum_{i=1}^{k}\frac{c_2\sqrt{ix}}{i^{3/2}} = \frac{c_1 c_2 x}{4}\sum_{i=1}^{k}\frac{1}{i} \sim \frac{c_1 c_2 x}{4}\log k,$$

which shows the contradiction if we take k large enough.

We can prove by similar methods that if $\liminf_{x\to\infty} B(x)/\sqrt{x} > 0$, then for every $\varepsilon > 0$ there is a $c > 0$ such that for infinitely many x

$$A(x(1+c)) - A(x) < \varepsilon\sqrt{x}. \tag{25}$$

Perhaps (25) can be replaced by

$$A\{A(x(1+c)) - A(x)\} + \{B(x(1+c)) - B(x)\} = o(\sqrt{x}). \tag{26}$$

At present we cannot prove (26).

References

1. M. Ajtai, J. Komlós, and E. Szemerédi, A dense infinite Sidon sequence, *European J. Combin.* **2** (1981), 1–11.
2. P. Erdős and R. L. Graham, "Old and New Problems and Results in Combinatorial Number Theory," Monographie No. 28 de L'Enseignement Mathématique, Genève, 1980.
3. H. Halberstam and K. F. Roth, "Sequences," Oxford Univ. Press, London/New York, 1966.