Europ. J. Combinatorics (1985) **6**, 201–203

An Application of Graph Theory to Additive Number Theory

NOGA ALON* AND P. ERDÖS

A sequence of integers $A=\{a_1<a_2<\cdots<a_n\}$ is a $B_2^{(k)}$ sequence if the number of representations of every integer as the sum of two distinct a_is is at most k. In this note we show that every $B_2^{(k)}$ sequence of n terms is a union of $c_2^{(k)}\cdot n^{1/3}$ $B_2^{(1)}$ sequences, and that there is a $B_2^{(k)}$ sequence of n terms which is not a union of $c_1^{(k)}\cdot n^{1/3}$ $B_2^{(1)}$ sequences. This solves a problem raised in [3, 4]. Our proof uses some results from extremal graph theory. We also discuss some related problems and results.

Sidon called a finite or infinite sequence of integers $A=\{a_1<a_2<\cdots\}$ a $B_2^{(k)}$ sequence if the number of representations of every integer as the sum of two distinct a_is is at most k. In particular he was interested in $B_2^{(1)}$, or, for short, B_2 sequences, i.e. the case where all the sums a_i+a_j are distinct.

Let f_n denote the maximal cardinality of a B_2 subsequence of $\{1, 2, \ldots, n\}$. Turán and Erdös proved [5]

$$n^{1/2}-O(n^{5/16})<f_n<n^{1/2}+O(n^{1/4}). \tag{1}$$

The lower bound of (1) was also proved by Chowla. Let H_n denote the largest r such that every sequence of n integers contains a B_2 subsequence of cardinality r. Komlós, Sulyok and Szemerédi [6] proved a general theorem which implies

$$H_n>c\cdot n^{1/2}, \tag{2}$$

where c is an absolute constant. By (1) $c\leqslant 1$, and maybe,

$$H_n=(1+o(1))n^{1/2}.$$

This does not seem to be easy to prove.

Let $H_n^{(k)}$ denote the largest r such that every $B_2^{(k)}$ sequence of n integers contains a B_2 subsequence of cardinality r. In [3] an infinite $B_2^{(2)}$ sequence which is not the union of a finite number of B_2 subsequences is constructed. A similar construction shows that there exists a $B_2^{(2)}$ sequence of n terms with no B_2 subsequence of cardinality $\geqslant c\cdot n^{2/3}$ (see [4]). Thus

$$(H_n^{(k)}\leqslant)H_n^{(2)}<c\cdot n^{2/3}. \tag{3}$$

In this note we prove

THEOREM 1. *Every $B_2^{(k)}$ sequence of n terms is a union of $c_2^{(k)}\cdot n^{1/3}$ B_2 sequences. On the other hand, by* (3) *there is a $B_2^{(k)}$ sequence of n terms which is not a union of $c_1^{(k)}\cdot n^{1/3}$ B_2 sequences.*

At the moment we cannot strengthen this result to $(c_3^{(k)}+o(1))n^{1/3}$. It is perhaps interesting to observe that the dependence on k is so weak. Note that Theorem 1 implies that

$$H_n^{(k)}\geqslant c_4^{(k)}\cdot n^{2/3}. \tag{4}$$

This solves a problem raised in [3, 4].

* Research supported in part by the Weizmann Fellowship for Scientific Research.

0195-6698/85/030201+03 $02.00/0

PROOF OF THEOREM 1. Since $(3/c)(n-c\cdot n^{2/3})^{1/3}+1\leq(3/c)n^{1/3}$, repeated application of (4) implies the assertion of Theorem 1 (with $c_2^{(k)}=3/c_4^{(k)}$). We thus have to prove (4). Let $A=\{a_1<a_2<\cdots<a_n\}$ be a $B_2^{(k)}$ sequence. Let $G=(V,E)$ be a 4-uniform hypergraph on the set of vertices $V=\{1,2,\ldots,n\}$ where $\{i,j,l,m\}$ is an edge if $a_i+a_j=a_l+a_m$. The number of edges of G is clearly $<\frac{1}{2}(k-1)\cdot\binom{n}{2}\leq\frac{1}{4}(k-1)\cdot n^2$. Note that if $F\subseteq V$ is independent, (i.e. no edge of G is contained in F), then $\{a_f; f\in F\}$ is a B_2 subsequence of A. Thus we have to show that G contains an independent subset of size $\geq c(k)\cdot n^{2/3}$. This follows either from the known results about Turán's problem for hypergraphs (see, e.g. D. de Caen [1, inequality (5)]) or from an easy application of the probabilistic method. Indeed, choose every vertex in V independently with probability $c\cdot n^{-1/3}$ to obtain a subset U of V of cardinality $(c+o(1))\cdot n^{2/3}$ containing $\leq((k-1)/4+o(1))c^4\cdot n^{2/3}$ edges. F is obtained from U by deleting one vertex from each such edge. If $c=c(k)$ is chosen appropriately we clearly obtain the desired result. This completes the proof.

Using a similar, though somewhat more complicated, probabilistic argument we can show that the analogue of (4) holds also for infinite sequences, namely:

THEOREM 2. *Every infinite $B_2^{(k)}$ sequence $A=\{a_1<a_2<\cdots\}$ contains a B_2 subsequence C such that for every $n\geq 1$*

$$|C\cap\{a_1,a_2,\ldots,a_n\}|\geq[c^{(k)}n^{2/3}]. \tag{5}$$

OUTLINE OF PROOF. For $i\geq 1$ choose, independently, a_i with probability $c/i^{1/3}$ to get a sequence $D=\{d_1<d_2<\cdots\}$. A quadruple $\{d_i,d_j,d_l,d_m\}$ of elements of D is *bad* if $d_i+d_j=d_l+d_m$. Let C be the subsequence of A obtained from D by deleting the largest element of every bad quadruple. Obviously D is a B_2 sequence.

Easy estimates of the expected values and the variances of the random variables $|D\cap\{a_1,\ldots,a_n\}|$ and $|\{Q\colon Q \text{ is a bad quadruple in } D\cap\{a_1,\ldots,a_n\}\}|$ show that if $c=c(k)$ is sufficiently small, then, with positive probability, (5) holds for all $n=2^r$. This implies the validity of (5) (with a smaller constant $c^{(k)}$) for all $n>0$.

Another property of $B_2^{(k)}$ sequences is given in the following theorem.

THEOREM 3. *Every (finite or infinite) $B_2^{(k)}$ sequence is a union of $c=c(k)$ subsequences, each of which contains no arithmetic progression of three terms.*

PROOF. Let $A=\{a_1<a_2<\cdots\}$ be a $B_2^{(k)}$ sequence. Let $G=(V,E)$ be a 3-uniform hypergraph on the set of vertices $V=\{1,2,\cdots\}$ in which $\{i,j,l\}$ is an edge if $a_i+a_j=2a_l$. We must show that V can be covered by $c(k)$ independent subsets. Let H be an induced subgraph of G on r vertices. Clearly H contains at most $r\cdot k$ edges and hence contains a vertex of degree at most $3k$. Thus, by an easy induction, the vertices of any finite subgraph of G can be partitioned to $\leq 3k+1$ independent subsets. This proves the theorem for finite sequences. The infinite case follows, by the compactness principle.

Similar to Theorem 1 is the following.

THEOREM 4. *Every $B_2^{(k)}$ sequence of n terms is a union of $c_2^{(k)}\cdot n^{1/(2k-1)}$ $B_2^{(k-1)}$ subsequences. On the other hand if $k=2^s$ there exists a $B_2^{(k)}$ sequence of n terms which is not the union of $c_1^{(k)}\cdot n^{1/(2k-1)}$ $B_2^{(k-1)}$* subsequences.

PROOF. The first part of the theorem is proved as before. For the second part, we consider the following construction. Put $n = m^{2k-1}$. Let $A_0, A_1, A_2, \dots, A_s$ be disjoint sets of integers, $|A_i| = m^{2^i}$. Let $G = (V, E)$ be the complete $(s+1)$-uniform $(s+1)$-partite hypergraph on the classes of vertices $A_0, \dots, A_s$, i.e. $V = \bigcup_{i=0}^{s} A_i$ and E consists of all $(s+1)$-subsets of V having exactly one element from each A_i. Clearly $|E| = \prod_{i=0}^{s} |A_i| = n$. For each edge $e \in E$, put $a_e = \sum_{v \in e} 10^v$. One can easily check that $A = \{a_e;\, e \in E\}$ is a $B_2^{(k)}$ sequence of n terms. A standard hypergraph theoretic argument (analogous to that of [2]) shows that every subgraph of G of more than $c(k)n^{1-1/(2k-1)} = c(k)m^{2k-2}$ edges contains a copy of a complete $(s+1)$-partite hypergraph with 2 vertices in each class. Therefore for every subsequence D of A of more than $c(k)n^{1-1/(2k-1)}$ terms there are $a_i^1, a_i^2 \in A_i$ $(0 \le i \le s)$ such that all the 2^{s+1} numbers $\sum_{i=0}^{s} 10^{a_i^{\varepsilon_i}}$ $(\varepsilon_i \in \{1, 2\})$ are in D, and hence D is not a $B_2^{(k-1)}$ sequence. Thus no $B_2^{(k-1)}$ subsequence of A has cardinality $> c(k)n^{1-1/(2k-1)}$ and the assertion of the theorem follows.

It seems likely that every sequence of n terms is a union of $(1+o(1))n^{1/2}$ B_2-subsequences, but this seems to be very difficult, (and would imply, of course, that $c = 1 + o(1)$ in (2)). However, one can easily modify the proof of the lower bound of (1) to show that $\{1, 2, \dots, n\}$ is a union of $(1+o(1))n^{1/2}$ B_2-sequences.

The method of this note implies easily that for every $\varepsilon > 0$ there exists a $c = c(\varepsilon)$ such that the sequence $\{1, 2^2, 3^2, 4^2, \dots, n^2\}$ contains a B_2-subsequence of cardinality $c \cdot n^{2/3-\varepsilon}$. We do not know how close this bound is to the truth. Maybe $n^{2/3-\varepsilon}$ can be replaced by $n^{1-\varepsilon}$. However, by Landau's well known result on the density of the sums of two squares one can easily show an upper bound of $c' \cdot n/(\log n)^{1/4}$ for this cardinality.

We conclude this note with another problem. Call an (infinite) sequence $\{a_1 < a_2 < \cdots\}$ free if for any two distinct sets of indices I, J $\sum_{i \in I} a_i \ne \sum_{j \in J} a_j$. Pisier was interested in a condition that guarantees that a sequence A is a union of a finite number of free subsequences. He observed that a necessary condition is:

There exists a $\delta > 0$ such that every finite subsequence B of A has a free subsequence C of cardinality $\ge \delta |B|$. (6)

It seems unlikely that (6) is also sufficient. However, we could not find any counterexample. One can formulate, of course, the analogous problem for B_2 sequences.

REFERENCES

1. D. de Caen, Extension of a theorem of Moon and Moser on complete subgraphs, *Ars Combinatoria* **16** (1983), 5–10.
2. P. Erdös, On extremal problems on graphs and generalized graphs, *Israel J. Math.* **2** (1964), 183–190.
3. P. Erdös, Some applications of Ramsey's Theorem to additive number theory, *Europ. J. Combinatorics* **1** (1980), 43–46.
4. P. Erdös, Extremal problems in number theory, combinatorics and geometry, *Proc. Inter. Congress in Warsaw*, 1983 (to appear).
5. Many references to B_2 sequences can be found in: H. Halberstam and K. F. Roth, *Sequences*, Clarendon Press, Oxford, 1966, chapters 2, 3.
6. J. Komlós, M. Sulyok and E. Szemerédi, Linear problems in combinatorial number theory, *Acta Math. Hungar. Acad. Sci.* **26** (1975), 113–121.

Received 20 May 1984

N. ALON
Department of Mathematics, Massachusetts Institute of Technology and Mathematics Research Centre, AT&T Bell Laboratories, Murray Hill, NJ 07974, U.S.A.
and
P. ERDÖS
Mathematical Institute, Hungarian Academy of Sciences Budapest H-1364, Hungary