J. Number Th., 56 (1996), 1, 4-11

The Density of $B_h[g]$ Sequences and the Minimum of Dense Cosine Sums

Mihail N. Kolountzakis
Department of Mathematics, Stanford University, Stanford, CA 94305
E-mail: kolount@math.stanford.edu

April 1993; revised May 1994

Abstract

A set E of integers is called a $B_h[g]$ set if every integer can be written in at most g different ways as a sum of h elements of E. We give an upper bound for the size of a $B_h[1]$ subset $\{n_1, \ldots, n_k\}$ of $\{1, \ldots, n\}$ whenever $h = 2m$ is an even integer:

$$k \le \left(m(m!)^2\right)^{1/h} n^{1/h} + O(n^{1/2h}).$$

For the case $h = 2$ ($h = 4$) this has already been proved by Erdős and Turán (by Lindström). It has been independently proved for all even h by Jia [9] who used an elementary combinatorial argument. Our method uses a result, which we prove, related to the minimum of dense cosine sums which roughly states that if $1 \le \lambda_1 < \cdots < \lambda_N \le (2-\epsilon)N$ are N different integers then

$$\left|\min_x \sum_1^N \cos \lambda_j x\right| \ge C\epsilon^2 N.$$

Finally we exhibit some dense finite and infinite $B_2[2]$ sequences.

Mathematics Subject Classification: Primary 11B13, 42A05

1 Introduction

Let E be a set of integers. For any integer x we denote by $r_E(x;h)$ the number of ways x can be written as a sum of h (not necessarily distinct) elements of E. Two sums $a_1 + \cdots + a_h$ and $b_1 + \cdots + b_h$ are considered the same if the a_j's are a permutation of the b_j's. A set E of integers is called a $B_h[g]$ set if $r_E(x;h) \le g$ for all x. A $B_h[1]$ set is also called a B_h set and B_2 sets are sometimes called Sidon sets (the term "Sidon set" appears with a different meaning in harmonic analysis). The letter C stands for an arbitrary positive constant in this paper.

We are interested here in the density of $B_h[g]$ sets. Considerably more is known about B_2 sets than general $B_h[g]$ sets ([8, Ch. 2] is the principal reference). The main reason for this is that a set E is B_2 if and only if all differences $x-y$, for $x,y\in E$, $x\neq y$, are distinct. Nothing similar is true for $B_2[g]$ sets, for example.

Let $F_h(n)$ be the maximum size of a B_h set contained in $\{1,\ldots,n\}$.

It is obvious that $F_h(n)\le A_h n^{1/h}$ and it is the size of the constant A_h that we care about in this paper. For $h=2$ Erdős and Turán [5], using a counting method, have proved $F_2(n)\le \sqrt{n}+O(n^{1/4})$. (The constant one obtains by just counting differences is $\sqrt{2}$.) For $h=4$ Lindström [14], using van der Corput's lemma, has proved $F_4(n)\le (8n)^{1/4}+O(n^{1/8})$. In the other direction it has been shown by Chowla and by Erdős (see references in [8, Ch. 2]) using a theorem of Singer [15] that $F_2(n)\ge\sqrt{n}-o(\sqrt{n})$ and more generally it has been proved by Bose and Chowla [1] that $F_h(n)\ge n^{1/h}-o(n^{1/h})$.

The Cosine Problem: Chowla [4] has conjectured that for any distinct positive integers $n_1<\cdots<n_N$

$$\left|\min_x \sum_{j=1}^N \cos n_j x\right| \ge C\sqrt{N}. \tag{1}$$

This conjecture has remained unproved and the best result known to date is due to Bourgain [2]: $\left|\min_x \sum_{j=1}^N \cos n_j x\right| \ge 2^{\log^\epsilon N}$ for some $\epsilon>0$.

It is easy to see that there are sequences $\{n_j\}$ for which the left hand side of (1) is bounded above by $C\sqrt{N}$. The author proved in [11] using a probabilistic method that there are very dense such sequences: we can have $n_N\le 2N$. (This can also be proved using the lower bound on $F_2(n)$ mentioned above.) In Section 2 we prove that the above density is best possible (Theorem 2).

In Section 3 we use this result to prove the following theorem.

Theorem 1 *Let $h=2m\ge 2$ be an even integer. Then*

$$F_h(n)\le \left(m(m!)^2\right)^{1/h} n^{1/h}+O(n^{1/2h}). \tag{2}$$

Theorem 1 contains the results of Erdős and Turán and of Lindström as special cases. It has recently been proved independently by Jia [9] who used an elementary combinatorial argument.

In Sections 4 and 5 we show that allowing $g>1$ indeed helps. We exhibit a $B_2[2]$ subset of $\{1,\ldots,n\}$ with $\sqrt{2n}+o(\sqrt{n})$ elements and an infinite $B_2[2]$ sequence $1\le n_1<\cdots<n_j<\cdots$ for which

$$\liminf_j \frac{n_j}{j^2}=1.$$

Jia [10] has improved the method of Section 4 giving a $B_2[2]$ subset of $\{1,\ldots,n\}$ with $\sqrt{3n}+o(\sqrt{n})$ elements, and has generalized the results of Sections 4 and 5 to $B_2[g]$ sequences for $g>2$.

Results for the case of odd h (B_{2m-1} sequences) have been obtained by Li [13] (for B_3 sets), Chen [3] and Graham [7]. Chen and Graham proved that

$$F_{2m-1}(n) \le (m!)^{2/(2m-1)} n^{1/(2m-1)} + O(n^{1/(4m-2)}).$$

2 Dense Cosine Sums

It was proved in [11] that for every positive integer N there are positive integers $1 \le \lambda_1 < \cdots < \lambda_N \le 2N$ such that $\left|\min_x \sum_1^N \cos \lambda_j x\right| \le C\sqrt{N}$. We now prove that we cannot have more dense cosine sums whose minimum is small in absolute value.

Theorem 2 *Let $0 \le f(x) = M + \sum_1^N \cos \lambda_j x$, with $1 \le \lambda_1 < \cdots < \lambda_N \le (2-\epsilon)N$, for some $\epsilon > 3/N$. Then*

$$M > C\epsilon^2 N. \tag{3}$$

Proof: We use the following theorem of Fejér [6]:
Let $p(x)$ be a nonnegative trigonometric polynomial of degree n and constant term $\widehat{p}(0) = 1$. Then $p(0) \le n+1$.
The obvious inequality above is $p(0) \le 2n+1$. We note that Fejér's theorem is a corollary of the well known theorem of Fejér and Riesz which states that every nonnegative trigonometric polynomial can be written as the square of the modulus of a polynomial of the same degree.

To use Fejér's theorem we first need to "smooth" $\widehat{f}$. Define $p(x) = f(x)K_a(x) \ge 0$, where

$$K_a(x) = \sum_{j=-a}^{a} \left(1 - \frac{|j|}{(a+1)}\right) e^{ijx} \ge 0$$

is the Fejér kernel of degree a (the parameter a will be determined later). Then

$$\begin{aligned} \deg p &\le (2-\epsilon)N + a, \\ \widehat{p}(0) &= M + \frac{1}{a+1} \sum_{\lambda_j \le a} (a+1-\lambda_j), \\ p(0) &= (M+N)(a+1). \end{aligned}$$

Observe that $\widehat{p}(0) \le M + a/2$ (the worst case clearly being $\lambda_j = j$) and apply Fejér's theorem in the form $\widehat{p}(0) \ge p(0)/(1+\deg p)$ to get

$$M + \frac{a}{2} \ge \frac{(M+N)a}{(2-\epsilon)N + a + 1} \ge \frac{Na}{(2-\epsilon)N + a + 1} \ge \frac{1}{2-\epsilon+(a+1)/N} a.$$

Move $a/2$ to the right hand side and let $a = \epsilon N/2$ to get

$$M \ge \left(\frac{1}{4-\epsilon+2/N} - \frac{1}{4}\right)\epsilon N \ge \frac{(\epsilon - 2/N)\epsilon}{16 - 4\epsilon + 8/N} N \ge \frac{(\epsilon-2/N)\epsilon}{16} N \ge \frac{\epsilon^2}{3 \cdot 16} N,$$

since $\epsilon > 3/N$ implies $4\epsilon > 8/N$ and $\epsilon - 2/N > \epsilon/3$. □

3 An Upper Bound for $F_h(n)$, h Even

Let $h = 2m \geq 2$ be a fixed even integer. We shall give an upper bound for the density of B_h sets contained in $\{1, \ldots, n\}$. Theorem 2 is the main tool. In this section C denotes an arbitrary positive constant which may depend on h only.

Let $E = \{n_1, \ldots, n_k\}$, $1 \leq n_1 < \cdots < n_k \leq n$, be a B_h set. This means that all sums $a_1 + \cdots + a_h$ with $a_j \leq a_{j+1}$ and $a_j \in E$ are distinct. Consequently the sums of the form

$$a_1 + \cdots + a_m - b_1 - \cdots - b_m$$

with

$$a_j, b_j \in E,\ a_j < a_{j+1},\ b_j < b_{j+1} \text{ and } a_i \neq b_j \tag{4}$$

are all different. Indeed, if $\sum a_j - \sum b_j = \sum a'_j - \sum b'_j$ we have $\sum a_j + \sum b'_j = \sum a'_j + \sum b_j$ and, since E is a B_h sequence, the collection of terms in the left hand side is the same as that in the right hand side. But the a_j's have been assumed different from the b_j's, so we must have $a_j = a'_j$ and similarly $b_j = b'_j$.

Define the nonnegative polynomial

$$\begin{aligned} f(x) &= \left|\sum_{j=1}^{k} e^{in_jx}\right|^h = \left(\sum_{j=1}^{k} e^{in_jx}\right)^m \left(\sum_{j=1}^{k} e^{-in_jx}\right)^m \\ &= r(x) + 2(m!)^2 \left(\sum_{a_j, b_j \text{ satisfy } (4)} \cos\left(\sum a_j - \sum b_j\right)x\right). \end{aligned}$$

The polynomial $r(x)$ consists of $O(k^{h-1})$ terms with coefficient 1 and thus, for some $C > 0$,

$$Ck^{h-1} + \sum_{a_j, b_j \text{ satisfy } (4)} \cos\left(\sum_{j=1}^{m} a_j - \sum_{j=1}^{m} b_j\right)x \geq 0.$$

Write $\lambda_1 < \cdots < \lambda_N$, $N = k^h/(2(m!)^2) - O(k^{h-1})$, for the positive sums of the form $\sum a_j - \sum b_j$ (they are all different). Using Theorem 2 we conclude that

$$mn \geq \lambda_N \geq \left(2 - Ck^{-1/2}\right)\left(\frac{1}{2(m!)^2}k^h - O(k^{h-1})\right). \tag{5}$$

This implies

$$k \leq (m(m!)^2)^{1/h} n^{1/h} + o(n^{1/h}).$$

The error term can also be bounded as follows. Write $k = C_1 n^{1/h} + R$, where $R \geq 0$ and $C_1 = (m(m!)^2)^{1/h}$. Then $n = ((k - R)/C_1)^h$, and substituting this in (5) and matching the second largest terms we get $R = O(k^{1/2}) = O(n^{1/2h})$, which concludes the proof of Theorem 1.

Theorem 1 improves the estimate one gets by just counting the λ_j's. Indeed, there are $N = k^h/(2(m!)^2) - O(k^{h-1})$ different λ_j's in $\{1, \ldots, mn\}$ so we only get $k \leq \left(2m(m!)^2\right)^{1/h} n^{1/h} + o(n^{1/h})$.

4 Dense Finite $B_2[2]$ Sequences

As mentioned in Section 1, for each n there is a B_2 sequence $1 \le n_1 < \cdots < n_k \le n$ with $k = \sqrt{n} + o(\sqrt{n})$. In this Section we show that if one allows up to 2 sums to coincide we can have denser sequences. We do this by interleaving two dense B_2 sequences.

Theorem 3 *For each n there is a $B_2[2]$ set $B \subseteq \{1, \ldots, n\}$ with $|B| = \sqrt{2n} + o(\sqrt{n})$.*

Proof: By the lower bound on $F_2(n)$ there is a B_2 set $A \subseteq \{1, \ldots, \lfloor n/2 \rfloor - 1\}$, with $|A| = \sqrt{n/2} + o(\sqrt{n})$. We shall show that the subset $B = 2A \cup (2A+1)$ of $\{1, \ldots, n\}$ is $B_2[2]$ which proves the theorem.

The proof is by contradiction. Assume that we have the non-trivial relations

$$x_1 + y_1 = x_2 + y_2 = x_3 + y_3, \tag{6}$$

with $x_j, y_j \in B$ and let $z = x_1 + y_1$. Look at $x_j + y_j \bmod 2$. There are three possible patterns: $0+0$, $1+1$ and $0+1$.

If z is even then only $0+0$ and $1+1$ may appear in (6) and we have either a relation of the pattern $0+0 = 0+0$ or a relation of the pattern $1+1 = 1+1$. Both cases contradict the fact that A is B_2, the first after just dividing by 2, the second after cancelling the remainders and then dividing by 2.

If z is odd then only the pattern $0+1$ appears in (6) which can be rewritten as

$$2a_1 + (2a_1' + 1) = 2a_2 + (2a_2' + 1) = 2a_3 + (2a_3' + 1) \tag{7}$$

with $a_j, a_j' \in A$. By canceling 1 and dividing by 2 we have

$$a_1 + a_1' = a_2 + a_2' = a_3 + a_3'.$$

But A is B_2 so for at least one of the above relations, say the first one, we have $a_1 = a_2$ and $a_1' = a_2'$ which contradicts the fact that the first relation in (7) is non-trivial. □

Jia [10] has improved and generalized Theorem 3. He has proved the existence of a $B_h[g]$ set $B \subseteq \{1, \ldots, n\}$ such that

$$|B| = (m(h,g))^{1-1/h} n^{1/h} + o(n^{1/h}), \tag{8}$$

where $m(h,g)$ is the largest integer m for which the equation $a = x_1 + \cdots + x_h$ has at most g solutions in $\mathbf{Z}_m$ (up to rearrangement) for each $a \in \mathbf{Z}_m$. For $h = g = 2$ the coefficient of the major term in (8) becomes $\sqrt{3}$, thus improving our result.

5 Infinite $B_2[2]$ Sequences with Large Upper Density

The situation is rather different for infinite $B_2[g]$ sequences. Erdős [16] has proved that there is no infinite B_2 sequence $\{n_j\}$ with $n_j = O(j^2)$. Infinity is not the problem here but the fact that we require $n_j \le Cj^2$ for all j and not just for the last one, as we did with finite sequences.

For the upper density of the sequence $\{n_j\}$ Erdős [16] proved that it is possible to have $\liminf_j n_j/j^2 \le 4$ and Krückeberg [12] later improved this to $\liminf_j n_j/j^2 \le 2$. It is still unknown whether the number 2 in the right hand side above can be reduced (it cannot be less than 1 by the upper bound on $F_2(n)$ of Erdős and Turán).

We now show that for a $B_2[2]$ infinite sequence this is possible.

Theorem 4 *There is a $B_2[2]$ sequence $\{n_j\}$ with $\liminf_j n_j/j^2 = 1$.*

Proof: The theorem will be proved if we show that any $B_2[2]$ sequence $1 \le n_1 < \cdots < n_k$ can be extended to a sequence $1 \le n_1 < \cdots < n_k < n_{k+1} < \cdots < n_l$, such that $n_l = l^2 + o(l^2)$.

Write $A = \{n_1, \ldots, n_k\}$ and $x = n_k$. Take $B \subseteq \{2x+1, \ldots, x^4\}$ to be a B_2 set with $|B| = x^2 + o(x^2)$. In what follows $a_j \in A$, $b_j \in B$ and $d_j \in D$ (to be defined below).

Consider the relations of the form

$$a_1 + b_1 = a_2 + b_2. \tag{9}$$

Such a relation may be written as $a_1 - a_2 = b_2 - b_1$. But B is a B_2 set, so all differences $b_2 - b_1$ are distinct, which implies that a pair $a_1, a_2 \in A$ may appear in (9) only once. Thus there are $O(k^2) = O(x)$ of these relations which may involve $O(x)$ elements of B. Let then

$$D = \{b \in B : b \text{ does not appear in any relation of the form (9)}\} \tag{10}$$

and $E = A \cup D$. Obviously $|E| = x^2 + o(x^2)$. We show that E is a $B_2[2]$ set.

First note that the relations of the form

$$a_1 + a_2 = a_3 + d_1 \quad \text{or} \quad a_1 + a_2 = d_1 + d_2$$

are not possible (the left hand side is too small) and A is itself $B_2[2]$. This proves $r_E(a_1+a_2;2) \le 2$ for all $a_1, a_2 \in A$.

It remains to be checked that $r_E(a_1 + d_1; 2) \le 2$ and $r_E(d_1 + d_2; 2) \le 2$. By passing from B to D we eliminated all relations of the form (9) and so the only remaining non-trivial relations that we have to check are of the form

$$a_1 + d_1 = d_2 + d_3. \tag{11}$$

These are indeed possible. Assume $y = a_1 + d_1 = d_2 + d_3$. We have to show that these are the only ways that y can be written as a sum of two elements of E. But this is obvious since $y = d_2' + d_3'$ is impossible (this would mean $d_2 + d_3 = d_2' + d_3'$ which contradicts D in B_2), $y = a_1' + a_2'$ is impossible because of size and $y = a_1' + d_1'$ would mean that $a_1' + d_1' = a_1 + d_1$ which we took care to eliminate in (10). □

Remark: Because of the result of Section 4 the previous theorem is not necessarily best possible.

Jia [10] has generalized our result to any $g \geq 2$ obtaining an infinite $B_2[g]$ sequence $\{n_j\}$ with $\liminf_j n_j/j^2 = 1/\sqrt{2g-3}$.

References

[1] R. C. Bose and S. Chowla, *Theorems in the additive theory of numbers*, Comment. Math. Helv., 37 (1962-63), 141-147.

[2] J. Bourgain, *Sur le minimum d'une somme de cosinus*, Acta Arith., 45 (1986), 381-389.

[3] S. Chen, *On the size of finite Sidon sequences*, preprint.

[4] S. Chowla, *Some applications of a method of A. Selberg*, J. Reine Angew. Math., 217 (1965), 128-132.

[5] P. Erdős and P. Turán, *On a problem of Sidon in additive number theory and some related problems*, J. London Math. Soc., 16 (1941), 212-215.

[6] L. Fejér, *Über trigonometrische Polynome*, J. Reine Angew. Math., 146 (1915), 53-82.

[7] S. W. Graham, *Upper bounds for Sidon sequences*, preprint.

[8] H. Halberstam and K. F. Roth, *Sequences*, Springer-Verlag, New York, 1983.

[9] X.-D. Jia, *On finite Sidon sequences*, J. Number Th. 44 (1993), 84-92.

[10] X.-D. Jia, *$B_h[g]$ sequences with large upper density*, preprint.

[11] M. N. Kolountzakis, *On nonnegative cosine polynomials with nonnegative, integral coefficients*, Proc. Amer. Math. Soc. 120 (1994), 157-163.

[12] F. Krückeberg, *B_2 Folgen und verwandte Zahlenfolgen*, J. Reine Angew. Math., 106 (1961), 53-60.

[13] A.-P. Li, *On B_3 sequences*, Acta Math. Sinica 34 (1991), 1, 67-71.

[14] B. Lindström, *A remark on B_4 sequences*, J. Comb. Theory, 7 (1969), 276-277.

[15] J. Singer, *A theorem in finite projective geometry and some applications to number theory*, Trans. Amer. Math. Soc., 43 (1938), 377-385.

[16] A. Stöhr, *Gelöste und ungelöste Fragen über Basen der natürlichen Zahlenreihe II*, J. Reine Angew. Math., 194 (1955), 111-140.