UPPER AND LOWER BOUNDS FOR FINITE $B_h[g]$ SEQUENCES

JAVIER CILLERUELO, IMRE Z. RUZSA AND CARLOS TRUJILLO

Abstract. We give a non trivial upper bound, $F_h(g, N)$, for the size of a $B_h[g]$ subset of $\{1, \ldots, N\}$ when $g > 1$. In particular, we prove

$$F_2(g, N) \leq 1.864(gN)^{1/2} + 1$$

On the other hand we exhibit $B_2[g]$ subsets of $\{1, \ldots, N\}$ with

$$\frac{g + \lfloor g/2 \rfloor}{\sqrt{g + 2\lfloor g/2 \rfloor}} N^{1/2} + o(N^{1/2})$$

elements.

1. Upper bounds

Let $h \geq 2, g \geq 1$ be integers. A subset $A$ of integers is called a $B_h[g]$-sequence if for every positive integer $m$, the equation

$$m = x_1 + \cdots + x_h, \quad x_1 \leq \cdots \leq x_h, \quad x_i \in A$$

has, at most, $g$ distinct solutions.

Let $F_h(g, N)$ denote the maximum size of a $B_h[g]$ sequence contained in $[1, N]$. If $A$ is a $B_h[g]$ subset of $\{1, \ldots, N\}$, then $\binom{|A| + h - 1}{h} \leq ghN$, which implies the trivial upper bound

$$F_h(g, N) \leq (ghh!)^{1/h} \tag{1.1}$$

For $g = 1, h = 2$, it is possible to take advantage of counting the differences $x_i - x_j$ instead of the sums $x_i + x_j$, because the differences are all distinct. In this way, P. Erdős and P. Turán [2] proved that $F_2(1, N) \leq N^{1/2} + O(N^{1/4})$, which is the best possible except for the estimate of error term.

For $h = 2m$, Jia [4] proved $F_{2m}(1; N) \leq (m(m!)^2)^{1/2m} N^{1/2m} + O(N^{1/4m})$. A similar upper bound for $F_{2m-1}(1, N)$ has been proved independently by S.Chen [1] and S.W.Graham [3]: $F_{2m-1}(1, N) \leq ((m!)^2)^{1/2m-1} N^{1/2m-1} + O(N^{1/4m-2})$.

However, for $g > 1$, the situation is completely different because the same difference can appear many times, and, for $g > 1$ nothing better than (1.1) is known. In this paper we improve this trivial upper bound.

Theorem 1.1.

$$F_2(g, N) \leq 1.864(gN)^{1/2} + 1$$

$$F_h(g, N) \leq \frac{1}{(1 + \cos^h(\pi/h))^{1/h}} (hh!gN)^{1/h}, \quad h > 2$$

1
Let $A \subset [1, N]$ be $B_h$-sequence. $|A| = k$. Put $f(t) = \sum_{a \in A} e^{int}$. We have $f(t)^h = \sum_{n=h}^{hN} r_h(n)e^{int}$ where $r_h(n) = \#\{n = a_1 + \cdots + a_i; a_i \in A\}$

$$f(t)^h = h!g \sum_{n=h}^{hN} e^{int} - \sum_{n=h}^{hN} (h!g - r_h(n))e^{int} = h!gp(t) - q(t)$$

Since $r_h(n) \leq h!g$, we have

$$\sum_{n=h}^{hN} |h!g - r_h(n)| = \sum_{n=h}^{hN} (h!g - r_h(n)) = (h(N - 1) + 1)h!g - \sum_{n=h}^{hN} r_h(n) = (h(N - 1) + 1)h!g - k^h$$

thus

$$|q(t)| \leq hh!gN - k^h$$

for every value of $t$.

$p(t)$ is just a geometrical series and we can express it as

$$p(t) = e^{hit} \frac{1 - e^{i(h(N-1)+1)t}}{1 - e^{it}}$$

if $0 < t < 2\pi$. We shall use only the property that at values of the form $t = jth$, $t_h = \frac{2\pi}{h(N-1)+1}$ with integer $j$, $1 \leq j \leq h(N-1)$, we have $p(t) = 0$, thus $f(t)^h = q(t)$. Consequently

$$|f(jth)| \leq (hh!gN - k^h)^{1/h}$$

for any integer $j$, $1 \leq j \leq h(N-1)$.

Since the midpoint of the interval $[1, N]$ is $(N+1)/2$, it will be useful to express $f$ as

$$f(t) = \exp\left(\frac{N+1}{2}it\right)f^*(t),$$

where

$$f^*(t) = \sum_{a \in A} \exp\left((a - \frac{N+1}{2})it\right).$$

Now we consider a function $F(x) = \sum_{j=1}^{h(N-1)} b_j \cos(jx)$ satisfying $F(x) \geq 1$ for $|x| \leq \pi/h$. We define $C_F = \sum |b_j|$.

We are looking for a lower and an upper bound for $Re\left(\sum_{j=1}^{h(N-1)} b_j f^*(jth)\right)$.

(1.2) \[
Re\left(\sum_{j=1}^{h(N-1)} b_j f^*(jth)\right) \leq \sum_{j=1}^{h(N-1)} |b_j| |f^*(jth)| = \sum_{j=1}^{h(N-1)} |b_j| |f(jth)| \leq \left(\sum_{j=1}^{h(N-1)} |b_j| \right)(hh!gN - k^h)^{1/h} = C_F(hh!gN - k^h)^{1/h}.
\]

On the other hand

(1.3) \[
Re\left(\sum_{j=1}^{h(N-1)} b_j f^*(jth)\right) = Re\left(\sum_{a \in A} \sum_{j=1}^{h(N-1)} b_j e^{i(a - \frac{N+1}{2})jt_h}\right) = \]
\[= \sum_{a \in A} \sum_{j=1}^{b_j \cos \left( \left( a - \frac{N-1}{2} \right) t_h \right)} = \sum_{a \in A} F \left( \left( a - \frac{N-1}{2} \right) t_h \right) \geq k,\]

because \(|(a - \frac{N-1}{2})t_h| \leq \pi/h\) for any integer \(a \in A\).

From (1.2) and (1.3) we have

\[|A| = k \leq \frac{1}{(1 + \frac{1}{C_F})^{1/h}} \cdot (hh!gN)^{1/h}.\]

For \(h > 2\), we take \(F(x) = \frac{1}{\cos(\pi/h)} \cos(x)\), with \(C_F = \frac{1}{\cos(\pi/h)}\) and this proves the theorem for \(h > 2\).

For \(h = 2\), we can take \(F(x) = 2 \cos(x) - \cos(2x), C_F = 3\), which gives \(|A| \leq \sqrt{6} \sqrt{gN}\), a nontrivial upper bound.

However, an infinite series gives a better result. Take the function

\[F(x) = \begin{cases} 
1, & |x| \leq \pi/2 \\
1 + \pi \cos(x), & \pi/2 < |x| \leq \pi.
\end{cases}\]

It is easy to see that

\[F(x) = \frac{\pi}{2} \cos(x) + 2 \sum_{n=2}^{\infty} \frac{\cos(\pi n/2)}{n^2 - 1} \cos(nx).\]

This series satisfies the following: \(F(x) = 1\) for \(|x| \leq \pi/2\) with

\[C_F = \pi/2 + 2 \sum_{n=1}^{\infty} \frac{1}{4n^2 - 1} = \pi/2 + 2 \sum_{n=1}^{\infty} \frac{1}{2n-1} \left( \frac{1}{2n-1} - \frac{1}{2n+1} \right) = \pi/2 + 1.\]

However we must truncate the series to the integers \(n \leq 2(N-1)\). Let

\[F_T(x) = \frac{\pi}{2} \cos(x) + 2 \sum_{n=2}^{2(N-1)} \frac{\cos(\pi n/2)}{n^2 - 1} \cos(nx).\]

Observe that

\[|F_T(x) - F(x)| \leq 2 \sum_{n=1}^{\infty} \frac{1}{n^2 - 1} = \frac{1}{2N-2}.\]

Now we consider the polynomial \(F^+(x)\) of degree \(N\). If \(|x| < \pi/2\) we have

\[|F^+(x)| = \frac{2N-2}{2N-3} |F(x) + F_T(x) - F(x)| \geq \frac{2N-2}{2N-3} (|F(x)| - |F_T(x) - F(x)|) \geq \frac{2N-2}{2N-3} \left( 1 - \frac{1}{2N-2} \right) = 1\]

and \(C_{F^+} \leq \frac{2N-2}{2N-3} (\pi/2 + 1)\).

Thus

\[|A| = k \leq \frac{2}{\left(1 + \frac{1}{C_{F^+}}\right) \cdot (gN)^{1/2}}.\]
A simple calculation gives
\[
\frac{2}{(1 + \frac{1}{C_2 F})^{1/2}} - \frac{2}{(1 + \frac{1}{C_2 F})^{1/2}} \leq \frac{1}{N}
\]
Then
\[
|A| = k \leq \frac{2}{(1 + \frac{1}{C_2 F})^{1/2}} (gN)^{1/2} + \sqrt{\frac{g}{N}} = \frac{2\pi + 4}{\sqrt{\pi^2 + 4\pi + 8}} \sqrt{gN} + \sqrt{\frac{g}{N}} \leq 1.864\sqrt{gN} + 1
\]
because, obviously, \( g \leq N \).

2. Lower bounds

Now we are interested in finite \( B_2[g] \) sequences as dense as possible. Kolountzakis [6] exhibits a \( B_2[2] \) subset of \( \{1, ..., N\} \) with \( \sqrt{2N^{1/2}} + o(N^{1/2}) \) elements taking \( A = (2A_0) \cup (2A_0 + 1) \) with \( A_0 \) a \( B_2[1] \) sequence contained in \( \{1, ..., [N/2]\} \).

In general it is easy to construct a \( B_2[g] \) subset of \( \{1, ..., N\} \) with \( (gN)^{1/2} + o(N^{1/2}) \) elements. In the sequel we improve these results

**Theorem 2.1.**

\[
F_2(g, n) \geq \frac{g + \lfloor g/2 \rfloor}{\sqrt{g + 2\lfloor g/2 \rfloor}} N^{1/2} + o(N^{1/2}).
\]

For \( g = 2 \) theorem 2 gives
\[
F_2(2, N) \geq \frac{3}{2} N^{1/2} + o(N^{1/2}).
\]

In general, for \( g \) even we get
\[
F_2(g, N) \geq \frac{3}{2\sqrt{2}} (gN)^{1/2} + o(N^{1/2}).
\]

And for \( g \) odd,
\[
F_2(g, N) \geq \frac{3 - (1/g)}{2\sqrt{2 - (1/g)}} (gN)^{1/2} + o(N^{1/2})
\]

**Remark.** Jia’s constructions of \( B_h(g) \) sequences in [5] does not work (Jia, personal communication). In the last step of the proof of theorem 3.1. of [5] we cannot deduce from the hypothesis that \( \{b_{s1}, \ldots, b_{sh}\} = \{b_{t1}, \ldots, b_{th}\} \). Jia’s argument can be modified if we define \( g_n(h, m) \) as the number of solutions of the equation \( a \equiv x_1 + \cdots + x_h \pmod{m} \). \( 0 \leq x_i \leq m - 1 \). It would imply the result \( |B| = \sqrt{gN} + o(\sqrt{N}) \). But for \( g = 2 \) it is the Kolountzakis’s construction [6].

We need some definitions and lemmas in order to construct \( B_2[g] \) sequences satisfying Theorem 2.1.

**Definition 2.1.** We say that \( a_0, a_1, ..., a_k \) satisfies the \( B^*[g] \) condition if the equation \( a_i + a_j = r \) has at most \( g \) solutions. (Here, \( a_i + a_j = a_j + a_i \) counts as two solutions if \( i \neq j \).)
Definition 2.2. We say that a sequence of integers $C$ is a $B_2 \pmod{m}$ sequence if $c_i + c_j \equiv c_k + c_l \pmod{m}$ implies \{c_i, c_j\} = \{c_k, c_l\}.

Lemma 2.2. If $a_0, a_1, \ldots, a_k$ satisfies the $B^*[g]$ condition, and $C$ is a $B_2 \pmod{m}$ sequence, then the sequence $B = \cup_{i=0}^{k}(C + ma_i)$ is a $B_2[g]$ sequence.

Proof. If $b_1 + b_1' = b_2 + b_2' = \ldots = b_{g+1} + b_{g+1}'$, $b_j, b_j' \in B$ we can write

$b_j = c_j + a_{j,m}$

$b_j' = c_j' + a_{j',m}$

$c_j, c_j' \in C$, $a_{j,i}, a_{j',i} \in \{a_0, \ldots, a_k\}$ where we have ordered the pairs $b_j, b_j'$ such that $c_j \leq c_j'$.

Then we have $c_j + c_j' \equiv c_k + c_k' \pmod{m}$ for all $j, k$, which implies $c_j = c_k$, $c_j' = c_k'$.

On the other hand, all the $g + 1$ sums $a_{j,i} + a_{j',i}$ are equal. Thus there exists $j, k$ such that $a_{i,j} = a_{i,k}$, $a_{i,j'} = a_{i,k'}$.

Then, for these $j, k$, we have $b_j = b_k$ and $b_j' = b_k'$.

□

Lemma 2.3. The subset $A^g = A_1^g \cup A_2^g = \{k; \ 0 \leq k \leq g - 1\} \cup \{g - 1 + 2k; \ 1 \leq k \leq \lfloor g/2 \rfloor\}$ satisfies the condition $B^*[g]$.

Proof. Let

$r(m) = \#\{a; \ a, m - a \in A^g\}$

$r_{ij}(m) = \#\{a; \ a \in A_i^g, m - a \in A_j^g\}, \ 1 \leq i, j \leq 2$

We have $r(m) = r_{11}(m) + 2r_{12}(m) + r_{22}(m)$, because $r_{12} = r_{21}$.

With this notation we will prove that $r(m) \leq g$ for any integer $m$. First we study the functions $r_{ij}$.

• $r_{11}(m)$

If $a, m - a \in A_1^g$, then $0 \leq a \leq g - 1$ and $0 \leq m - a \leq g - 1$, which implies

$max\{0, m - g + 1\} \leq a \leq min\{g - 1, m\}$.

Then

$r_{11}(m) = max\{0, min\{g - 1, m\} - max\{0, m - g + 1\} + 1\},$

and

$r_{11}(m) = \begin{cases} 
  m + 1, & 0 \leq m \leq g - 1 \\
  2g - m - 1, & g \leq m \leq 2g - 1 \\
  0, & 2g - 1 \leq m
\end{cases}$

• $r_{12}(m)$

If $a \in A_2^g$, $m - a \in A_1^g$, then $a = g - 1 + 2k, 1 \leq k \leq \lfloor g/2 \rfloor$ and

$0 \leq m - (g - 1 + 2k) \leq g - 1$, which implies

$max\{1, \frac{m - 2g + 2}{2}\} \leq k \leq min\{\lfloor g/2 \rfloor, \frac{m - g + 1}{2}\}$.

Since the $k$’s are integers, we can write

$max\{1, \frac{m - 2g + 3}{2}\} \leq k \leq min\{\lfloor g/2 \rfloor, \frac{m - g + 1}{2}\}.$
Then

\[
    r_{12}(m) = \begin{cases} 
        0, & m \leq g \\
        \lfloor \frac{m-g+1}{2} \rfloor, & g \leq m \leq 2g-1 \\
        \lfloor \frac{g}{2} \rfloor - \lfloor \frac{m-2g+1}{2} \rfloor, & 2g \leq m \leq 3g-1 \\
        0, & 3g-1 \leq m 
\end{cases}
\]

- \( r_{22}(m) \)

Obviously, if \( m \) is odd then \( r_{22}(m) = 0 \).

If \( a, m - a \in A_2 \), then \( a = g - 1 + 2k, m - a = g - 1 + 2j, 1 \leq j, k \leq \lfloor g/2 \rfloor \)
we have

\[
    1 \leq j = m/2 - (g - 1) - k \leq \lfloor g/2 \rfloor,
\]

which implies, if \( m \) is even, that

\[
    \max\{1, m/2 - g - \lfloor g/2 \rfloor + 1\} \leq k \leq \min\{m/2 - g, \lfloor g/2 \rfloor\}.
\]

Then

\[
    r_{22}(m) = \max\{0, \min\{m/2 - g, \lfloor g/2 \rfloor\} - \max\{1, m/2 - g - \lfloor g/2 \rfloor + 1\} + 1\}
\]

Therefore, if \( m \) is even

\[
    r_{22}(m) = \begin{cases} 
        0, & m < 2g \\
        m/2 - g, & 2g \leq m \leq 3g - 1 \\
        g + 2\lfloor g/2 \rfloor - m/2, & 3g \leq m \leq 4g - 2 \\
        0, & 4g - 2 < m 
\end{cases}
\]

Now, we are ready to calculate \( r(m) \).

- \( m \leq g - 1 \)

\[
    r(m) = r_{11}(m) = m + 1 \leq g
\]

- \( g \leq m \leq 2g - 1 \)

\[
    r(m) = r_{11}(m) + 2r_{12}(m) = 2g - m - 1 + 2\lfloor \frac{m-g+1}{2} \rfloor \leq 2g - m - 1 + m - g + 1 = g
\]

\( 2g \leq m \leq 3g - 1 \)

If \( m \) is odd, \( r(m) = 2r_{12}(m) = 2\lfloor g/2 \rfloor - \lfloor \frac{m-2g+1}{2} \rfloor \leq g \). If \( m \) is even, \( r(m) = 2r_{21}(m) + r_{22}(m) = 2\lfloor g/2 \rfloor - \lfloor \frac{m-2g+1}{2} \rfloor + m/2 - g = 2\lfloor g/2 \rfloor - (m - 2g) + m/2 - g = 2\lfloor g/2 \rfloor + g - m/2 \leq 2\lfloor g/2 \rfloor + g - (2g)/2 \leq g \).

\( 3g \leq m \leq 4g - 2 \)

If \( m \) is odd, \( r(m) = 0 \) If \( m \) is even, \( r(m) = r_{22}(m) = g + 2\lfloor g/2 \rfloor - m/2 \leq g + 2\lfloor g/2 \rfloor - (3g)/2 \leq g/2 < g \).

\( \square \)

Proof. (Theorem 2.1)

It is known [2], that for \( m = p^2 + p + 1 \), \( p \) prime, there exists a \( B_2 \) (mod \( m \))
sequence \( C_m \) such that \( |C_m| = p + 1 \) and \( C_m \subset [1, m] \)

Let us take

\[
    B = \cup_{i=0}^{k} (C_m + ma_i)
\]

where \( A^g = \{a_0, a_1, \ldots, a_k\} \) is defined in lemma 2.2.

Observe that \( B \subset [1, m(1 + a_k)] \), where \( a_k = g - 1 + 2\lfloor g/2 \rfloor \). Observe, also, that
\[
    |B| = |A^g||C_m| = (g + \lfloor g/2 \rfloor)(p + 1).
\]

Then \( F_p(g, m(2\lfloor g/2 \rfloor)) \geq (g + \lfloor g/2 \rfloor)(p + 1) \).

For any integer \( n \) we can choose a prime \( p \) such that

\[
    n - o(n) \leq (p^2 + p + 1)(g + 2\lfloor g/2 \rfloor) \leq n
\]
Then
\[ F_2[g, n] \geq F_2[g, m(g + 2\lceil g/2 \rceil)] \geq (g + \lfloor g/2 \rfloor)(p + 1) \geq \frac{g + \lceil g/2 \rceil}{\sqrt{g + 2\lceil g/2 \rceil}} n^{1/2} + o(n^{1/2}) \]

\[ \square \]

REFERENCES


[6]. M.N.Kolountzakis, “The density of \(B_h[g]\) sequences and the minimum of dense cosine sums”. *Journal of Number Theory* 56, 4-11 1996.