RECENT PROGRESS ON FINITE $B_h[g]$ SETS

ALAIN PLAGNE

1. Introduction

At the very beginning, there was Simon Sidon. This Hungarian mathematician was born in 1892. According to Babai [1], he was employed by an insurance company. He passed away in 1941.

The history of $B_h[g]$ sets begins in 1932, when Sidon, motivated by considerations on Fourier analysis [35], asks for the first time how large can be a sequence of distinct integers from $\{1, \dots, N\}$ with the property that all sums $a + b$, $a, b \in \mathcal{A}$, $a \leq b$ be different. Sets of integers with this property are now called Sidon sets, B_2 or $B_2[1]$ sets. The link between this question and harmonic analysis is via the study of (lacunary) power series $\sum_{i>0} z^{n_i}$, the h-th power of which has bounded coefficients.

More generally, a set of integers $\mathcal{A}$ is said to belong to the class $B_h[g]$ if for any integer n, the following upper bound holds:

$$\left|\{(a_1, \dots, a_h) \in \mathcal{A}^h : a_1 + \cdots + a_h = n, a_1 \leq \cdots \leq a_h\}\right| \leq g. \tag{1}$$

To avoid trivialities, we naturally have to impose the conditions $h \geq 2, g \geq 1$. The maximal cardinality of a $B_h[g]$ set of integers taken from $\{1, \dots, N\}$ will be denoted by $F_{h,g}(N)$. The asymptotic behavior of this function ($N \to +\infty$) is the central topic of the theory and it is our aim here to present several attempts to approach $F_{h,g}(N)$. In other words, we present here an historical account on the subject with some new results and a special emphasis on methods and perspectives. Several open questions are also stated.

It is fair and natural to quote previous general works on the subject as the nice book [17] (Chapter 2), the electronic survey [23] or the paper [34]. The reader who is lucky enough to have someone with Hungarian as his native language in his neighborhood is referred to [11]. I enjoyed Zoltán Szigeti's help in translating large parts of this popularization paper.

A related, although rather different, problem that we do not discuss here is that of *infinite* $B_h[g]$ sets. Specifically, the main problem is how dense such a set can be? Even for the original Sidon sets this question is not solved. If $F_{\mathcal{S}}$ denotes the counting function of the Sidon set $\mathcal{S}$, it is not known how large

$$\alpha = \sup \left\{ \liminf_{N \to +\infty} \frac{\log F_{\mathcal{S}}(N)}{\log N}, \mathcal{S} \text{ Sidon set} \right\}$$

is. The current state of knowledge on this problem is only $\sqrt{2} - 1 \leq \alpha \leq 1/2$. The upper bound follows from the finite case (see relation (2) below) while the lower bound has been proved by Ruzsa in a beautiful paper [32]. Anyway the gap between these two bounds remains large. Another related problem is to study $B_h[g]$ sets whose elements are not integers but elements of a given group [2]: modular Sidon sets are of special interest especially in the context of *difference-set* theory, a theory born at the same time as that of $B_h[g]$ sets in [36] but continued in [18] (see [21]).

The historical question which dealt only with $B_2[1]$ sets has been answered rather quickly in the forties by Erdős, Túran and Singer [12, 36] and the answer is

$$F_{2,1}(N) \sim \sqrt{N}. \tag{2}$$

Concerning the general question of $B_h[g]$ sets, lower bounds have been obtained by Bose and Chowla [4] some years later but, concerning the upper bounds, nothing else than the trivial counting argument - let $\mathcal{A} \subset \{1, \ldots, N\}$ be a $B_h[g]$ set, there are $\sim \binom{|\mathcal{A}|}{h}$ ordered h-tuples and all the corresponding sums are less than hN, each value being taken at most g times - which furnishes

$$F_{h,g}(N) \lesssim (ghh!)^{1/h} N^{1/h} \tag{3}$$

was known at the beginning of the sixties. Here and in the sequel, the notation $f(N) \lesssim g(N)$ means $f(N) \leq (1 + o(1))g(N)$ as N tends to infinity. In [17] their celebrated book *Sequences* published in 1966, Heini Halberstam and Klaus Friedrich Roth assert on page 96:

> We have seen that substantial progress has been made with questions concerning $B_2[1]$ sequences.(...) In comparison, very little is known concerning the corresponding questions relating to $B_2[g]$ sets.(...) It would be desirable to obtain corresponding results (...).

Since then several developments have been given but still there is no satisfactory answer to the question of the behavior of $F_{h,g}(N)$. It is not even known whether or not $F_{h,g}(N)N^{-1/h}$ converges. In this article, we report on the most recent advances in the theory and on some specific questions. We underline the fact that rather different techniques (purely combinatorial, analytical, probabilistic, etc.) can be applied and that perhaps an approach mixing together all these techniques is desirable.

To conclude, we take the opportunity of this introduction to stress the fact that there exist other Sidon sets. Unfortunately (but this is quite natural since Sidon was mostly interested in the subject), the other Sidon sets are also related with harmonic analysis. The interested reader can refer to [28].

2. Lower bounds

2.1. **Early results.** Beside the very first but rough estimate by Sidon himself (namely, according to [12], $F_{2,1}(N) \gg N^{1/4}$) and the estimate $F_{2,1}(N) \gg N^{1/3}$, pointed out later by Mian and Chowla [29] which comes from the greedy algorithm (namely consider the sequence $(a_n)_{n>0}$ defined by $a_1 = 1$ and

$$a_{m+1} = \min\{x \in \mathbb{N} \text{ with } x > a_m, x \neq a_r + a_s - a_t, 1 \leq r, s, t \leq m\}$$

for $m \geq 1$), the first explicit result on lower bounds was given by Erdős and Túran in their 1941 paper [12]: they proved the existence of a Sidon set in $\{1, \dots, N\}$ with more than $(1/\sqrt{2} - \varepsilon)\sqrt{N}$ elements for any positive ε. In view of (3), this has already the right order of magnitude $\sqrt{N}$, but this is not optimal (by (2)). Three years later, Erdős [10] realized that in 1938, Singer [36] proved the statement that if m is a power of a prime, there is a set of $m + 1$ elements which form a Sidon set modulo $m^2 + m + 1$ and consequently a Sidon set of integers. This now gives the exact asymptotic size of a maximal $B_2[1]$ sets (see formula (2)).

An essential fact for getting lower bounds is that it is sufficient to get good lower bounds for N prime. This is due to the fact that the ratio of two consecutive primes is asymptotically 1 (this follows for example from the prime number theorem) and also to the increasing behavior of $F_{h,g}(N)$ (with respect to N).

Bose [3] obtained a construction, similar to that of Singer, with m elements modulo m^2-1. In [33], Ruzsa proposed another approach which gives, for a prime p, a collection of $p-1$ elements that form a Sidon set modulo p^2-p; it is based on the existence of a primitive root modulo p.

Concerning $B_h[1]$ sets, Bose and Chowla obtained in [4] a generalization of Bose's result [3] and showed that if m is a power of a prime and $h \geq 2$, there is a set of m elements which form a $B_h[1]$ set modulo m^h-1.

All these results share in common that they are algebraic in nature and in fact rely on special constructions in finite fields (which thus needs essentially that the moduli used be prime - or prime powers).

These theorems imply $F_{h,1}(N) \gtrsim N^{1/h}$ and since a $B_h[g]$ set is a $B_h[g']$ set for $g' \geq g$, we get for all $h \geq 2, g \geq 1$,

$$F_{h,g}(N) \gtrsim N^{1/h}. \tag{4}$$

The main drawback of this formula is that is does not depend on g, which should be the case for a sharp formula. Anyway, this with (3) gives already the good order of magnitude $N^{1/h}$.

When $g=1$, formula (4) is still the best that is known. Thus, we conclude this section with the following question.

Problem 1. *For $h \geq 3$, find denser $B_h[1]$ sets (in other words improve on Bose-Chowla's construction) or show that the estimates of this section are best possible.*

According to the remark following the proof of Theorem 3.1 in [11], Bose-Chowla's construction could be nearly optimal (in other words, the formula $F_{h,1}(N) \sim N^{1/h}$ is plausible).

2.2. **More recent results.** Jia's paper [20] presents constructions of large $B_2[g]$ sets. Unfortunately, the argument is not complete and Jia's result must be weakened to

$$F_{2,g}(N) \gtrsim \sqrt{gN}, \tag{5}$$

a result already obtained, in the case $g = 2$, in [22] and even earlier (but in Hungarian!) in [11] as Theorem 3.13 (3.13 Tétel). Jia's result has been improved to

$$F_{2,g}(N) \gtrsim \frac{g + [g/2]}{\sqrt{g + 2[g/2]}} \sqrt{N} \tag{6}$$

by Cilleruelo, Ruzsa and Trujillo [9]. Here, as everywhere else, the notation $[x]$ is for the integral part of the real x.

Recently Lindström [27] proposed another generalization which yields

$$\frac{F_{h,g}(N)}{N^{1/h}} \gtrsim [g^{1/(h-1)}]^{1-1/h}. \tag{7}$$

When $h = 2$, (7) leads to Jia's formula and is thus weaker than (6). It has the advantage to be valid for any h and g (although much precise when g is a $(h-1)$-th power). It must be remarked that (7) together with (3) shows that for h fixed the dependence of $F_{h,g}(N)N^{-1/h}$ (N tending to infinity) is in $g^{1/h}$.

In [8], other new lower bounds are obtained by generalizing the methods of [9]: for $\varepsilon > 0, h \geq 3$ and g large enough (depending on h and ε) we have

$$\frac{F_{h,g}(N)}{N^{1/h}} \gtrsim (1-\varepsilon) \left(\frac{\pi g^2 h}{6}\right)^{1/2h}. \tag{8}$$

Although seemingly ineffective, this result can lead, after some efforts, to practical bounds: for instance, in the special case of $B_3[g]$ sets, what can be obtained is

$$\frac{F_{3,g}(N)}{N^{1/3}} \geq [\sqrt{4g/3}]^{2/3}. \tag{9}$$

All the preceding bounds [11, 22, 9, 8, 27] rely on the very same principle. Take $\{a_0 = 0, \ldots, a_k\}$ a $B_h[g]$ set and C any $B_h[1]$ set modulo m, then $\cup_{i=0}^{k}(C + ma_i)$ is a $B_h[gh!]$ set of integers. We thus get the estimate

$$\frac{F_{h,gh!}(N)}{N^{1/h}} \gtrsim \frac{k+1}{(1+a_k)^{1/h}}. \tag{10}$$

At this point, we add an important remark: the factor $h!$ appearing here is due to the possible permutations of the elements composing a sum $a_{i_1} + a_{i_2} + \cdots + a_{i_h}$. But for instance if all the elements in such a sum are equal, all these permutations are trivial. This shows that the truly interesting objects are the $B_h^*[g]$ sets, which are defined as those sets of integers such that, for any integer n,

$$\left|\{(a_1,\dots,a_h)\in\mathcal{A}^h : a_1+\cdots+a_h=n\}\right| \le g.$$

With respect to (1), the ordering of summands is removed. The important fact to see is that a $B_h[g]$ set is also a $B_h^*[gh!]$ set but there are $B_h^*[g]$ set which are not a $B_h[g']$ set with $g'h! = g$.

In the above-mentioned papers, a bound is obtained with general forms of $B_h[g]$ or $B_h^*[g]$ sets, that is the set $\{a_0,\dots,a_k\}$ is composed, for instance, with the first integers ($a_i = i$) or with two sequences of consecutive integers. The advantage is that computations are more general and easier to perform, the drawback is that it is not optimal, generally speaking.

We now present a way to obtain improvements on the known lower bounds. Our approach generalizes that of our paper [15]. Define

$$\mathcal{M}_{h,g} = \left\{ \frac{k+1}{(1+a_k)^{1/h}} \text{ where } k\ge 0, \{a_0,\dots,a_k\}\in B_h^*[g] \right\}$$

and $\mu_{h,g} = \sup \mathcal{M}_{h,g}$. Whenever one proves that $x \in \mathcal{M}_{h,g}$, it follows $F_{h,g}(N)N^{-1/h} \gtrsim x$ and thus

$$\frac{F_{h,g}(N)}{N^{1/h}} \gtrsim \mu_{h,g}. \tag{11}$$

Clearly we are interested in finding $B_h^*[g]$ sets with a corresponding ratio (density) $(k+1)(1+a_k)^{-1/h}$ large. We underline the fact that this quantity is related to the quantity we study in this paper (but here we are looking for a supremum instead of a limit superior). It is natural to conjecture that the supremum of $\mathcal{M}_{h,g}$ is attained (see also Problem 6 below) for a relatively small set. Thus we are led to the following problem.

Problem 2. *Given g and h, show that $\mu_{h,g}$ is attained, identify the set which reaches this supremum and find the value of $\mu_{h,g}$ (or at least its asymptotic behavior when g and h are large).*

Even if our knowledge on $\mu_{h,g}$ is poor, the challenge is now to find lower bound for it as large as possible (hopefully optimally large) by considering small $B_h^*[g]$ sets. As simple as this technique is, it seems not to have been noticed since then except in [15] where the technique described above has been applied to the special case of $B_2[2]$ sets by L. Habsieger together with the author of these lines. We were able to answer Problem 2 in this case: $\mu_{2,2}$ is attained for the Sidon set $\{0,1,4,6\}$ which gives $\mu_{2,2} = 4/\sqrt{7}$. This leads to

$$F_{2,2}(N) \gtrsim \frac{4}{\sqrt{7}}\sqrt{N}, \tag{12}$$

a small improvement on the 3/2 that (6) gives since $4/\sqrt{7} = 1.5118\ldots$.

Let us give some other new lower bounds (in fact I am convinced that the technique would easily improve on almost any case to which it applies). Take the case of $B_2[4]$ sets, by (6) we obtain the value $3/\sqrt{2} = 2.1213\ldots$. Now consider the $B_2[2]$ (and consequently $B_2^*[4]$) set $\{0,1,3,9,10,11,13,18,24,25,29,30\}$. It yields

$$\mu_{2,4} \geq \frac{12}{\sqrt{31}} = 2.1552\ldots$$

a lower bound which I conjecture to be an equality. Thus applying (10) we obtain

$$\frac{F_{2,4}(N)}{\sqrt{N}} \gtrsim 2.1552\ldots$$

a simple lower bound which, as far as I am aware, was not known.

We can also show an improvement for $B_2[6]$ sets. Consider the set $\{0,1,2,3,4,5,8,9,12,14,18,19\}$. The reader (or his computer) can easily check that it is a $B_2[3]$ (and consequently $B_2^*[6]$) set. This gives

$$\mu_{2,6} \geq \frac{12}{\sqrt{20}} = 2.6832\ldots$$

and

$$\frac{F_{2,6}(N)}{\sqrt{N}} \gtrsim 2.6832\ldots$$

whereas (6) gives only 2.5980 ...

Let us now give an example of improvement on previous formulae using essentially $B_h^*[g]$ sets. By considering the $B_3^*[6]$ set $\{0,1,5,11,13\}$ (notice that this is not a $B_3[1]$ set since $5+5+1 = 0+0+11$), we obtain $\mu_{3,6} \geq \frac{5}{14^{1/3}}$ and

$$\frac{F_{3,6}(N)}{N^{1/3}} \gtrsim \frac{5}{14^{1/3}} = 2.0745\ldots,$$

which improves the value $2^{2/3} = 1.5874\ldots$ given by (7) or (9).

As our last example of new lower bound consider $\{0,1,3,4\}$; this is a $B_3^*[9]$ set. We thus obtain $\mu_{3,9} \geq \frac{4}{5^{1/3}}$ and

$$\frac{F_{3,9}(N)}{N^{1/3}} \geq \frac{4}{5^{1/3}} = 2.3392\ldots,$$

which improves the value $3^{2/3} = 2.0800\ldots$ given by (7) or (9).

A few other practical cases can be handled because of the difficulties in computing small good $B_h^*[g]$ sets. But no doubt that the procedure described here gives very often (slight) improvements on the best known results as soon as one knows what is the optimal set to take.

Anyway, from all of this, something appears clearly : to be more explicit, take the case of $B_2[2]$ sets. The lower bound we have obtained is $4/\sqrt{7} = 1.51\ldots$ but a reasonable conjecture states that the truth is about 2 (the best upper bound is about 2.3). This shows that our way to build lower bounds is far from satisfactory. This is not altogether astonishing since every result stated in this section is obtained by some kind of gluing process and thus is rather artificial. In fact, there is no reason why a big $B_h[g]$ set should be obtained with this kind of gluing process ... and we could even say that there are good reasons why it should precisely *not* be of this form. But for the moment there is no alternative to these types of lower bounds. In particular, we do not know any direct approach to construct $B_h[g]$ sets for $g > 1$. The algebraic constructions seem not easy to generalize. Clearly an idea is required here, in order to solve the following problem.

Problem 3. *Find denser $B_h[g]$ sets by avoiding completely the "gluing" process. In other words, find direct constructions (like Bose-Chowla's for instance) of dense $B_h[g]$ sets.*

3. Upper bounds

The basic result concerning upper bounds is (3) which follows from a simple counting argument (Hajela's result [16] gives no improvement on it). This is not optimal as can be seen from the example of Sidon sets for which it is known (this is Lindström's form [25] of Erdős-Túran result) that

$$F_{2,1}(N) \leq \sqrt{N} + N^{1/4} + 1. \tag{13}$$

For the proof of this, the argument of [12] relies on counting small differences instead of sums (if all sums are distinct, clearly so are all non-zero differences). Although apparently insignificant, this simple remark is the key fact on which many papers on the subject rest. It is important to insist on this point in view of all the implications this remark has. To convince the reader, we just take the simplest possible example. Take a Sidon set $\mathcal{A} \subset \{1, \dots, N\}$. Using the fact that all the sums are different gives readily (3) that is

$$F_{2,1}(N) \lesssim 2N^{1/2}. \tag{14}$$

Now if $a, b \in \mathcal{A}$, $a < b$, $b - a$ is positive and different from all other differences. But it is less than N. Since there are $\sim |\mathcal{A}|^2/2$ such differences, we get $|\mathcal{A}|^2/2 \lesssim N$ or equivalently

$$F_{2,1}(N) \lesssim \sqrt{2}N^{1/2},$$

an improvement on (14) even though seemingly we did not do anything!

3.1. **The case** $g = 1$**.** When $h = 2$, we are led to the primitive Sidon sets. As seen above, this case is solved (and this is the only one) since (2) is known. Of course, we could now ask for precision on (13). We take this opportunity to quote the following problem of Erdős.

Problem 4. *Can (13) be improved (asymptotically)? Is* $F_{2,1}(N) \le \sqrt{N} + O(1)$ *true (Erdős conjecture) ? Or* $F_{2,1}(N) \le \sqrt{N} + O(N^{\varepsilon})$ *for any* $\varepsilon > 0$ *? More modestly, can one improve on the exponent* $1/4$ *in (13) ? At least, if we define*

$$\lambda = \limsup_{N \to +\infty} \frac{F_{2,1}(N) - \sqrt{N}}{N^{1/4}},$$

can one improve one $\lambda \le 1$*?*

Probably $\lambda < 1$ can be achieved, but what about proving $\lambda = 0$ (if true) ?

In the general case, the argument used by Erdős and Túran on small differences, can be partially reused (originally under the mask of a lemma by van der Corput) and an improvement on (3) was obtained first in the case $h = 4$ by Lindström [26]

$$F_{4,1}(N) \lesssim 8^{1/4} N^{1/4} = 1.6817 \dots N^{1/4}$$

and next for $h = 3$ by Li [24]

$$F_{3,1}(N) \lesssim 4^{1/3} N^{1/3} = 1.5874\ldots N^{1/3},$$

which was slightly improved by Graham [13] to $1.5868\ldots N^{1/3}$. In general, what was obtained is

$$F_{h,1}(N)N^{-1/h} \lesssim \begin{cases} \left(\frac{h}{2}\left(\frac{h}{2}!\right)^2\right)^{1/h} & \text{if } h \text{ is even,} \\ \left(\left(\frac{h+1}{2}\right)!\right)^{2/h} & \text{if } h \text{ is odd,} \end{cases}$$

the even case being from [19] (by a combinatorial method) and independently from [22] (by an analytical method) and the odd case in [5] (see also [13]). Combinatorial, analytical and probabilistic ideas led Cilleruelo, after some pioneering work in this area by Alon (see [23]), to improve on all these results in a recent preprint [7]: for instance the upper bounds obtained for $B_3[1]$ sets was about $1.5754N^{1/3}$ and that for $B_4[1]$ sets about $1.6739N^{1/4}$.

A breakthrough was recently introduced by Green [14]. It relies on Fourier analysis in $\mathbb{Z}/n\mathbb{Z}$ for some n depending on N. Standard tools such as convolution, Parseval's identity, Hölder inequality are then used. In the proof, an auxiliary function p is introduced so that at the end there is a numerical part for optimizing a quantity related to p. This is the flavor of the proof (we are aware of the frustrating character of this description of such a nice paper that we encourage the reader to refer to). This method gives the following results:

$$F_{3,1}(N) \lesssim \left(\frac{7}{2}\right)^{1/3} N^{1/3} = 1.5182\ldots N^{1/3},$$

$$F_{4,1}(N) \lesssim 7^{1/4} N^{1/4} = 1.6265\ldots N^{1/4}$$

and more generally

$$F_{h,1}(N)N^{-1/h} \lesssim \begin{cases} \left(\sqrt{\pi\frac{h}{2}}\left(\frac{h}{2}!\right)^2\right)^{1/h} & \text{if } h \text{ is even,} \\ \left(\sqrt{\frac{2\pi}{h+1}}\left(\frac{h+1}{2}!\right)^2\right)^{1/h} & \text{if } h \text{ is odd.} \end{cases}$$

We underline the fact that reappearance of techniques from harmonic analysis in Sidon's problem is all but unexpected in view of the area where the problem comes from.

3.2. **The case $h = 2$.** Cilleruelo, Ruzsa and Trujillo [9] obtained by another Fourier argument that

$$\frac{F_{2,g}(N)}{\sqrt{N}} \lesssim 1.864\sqrt{g}, \tag{15}$$

whereas formula (3) gives $2\sqrt{g}$.

The special case $B_2[2]$, which is the first non-solved case has attracted a big amount of efforts. In the paper [6], Cilleruelo proved by a combinatorial argument based on estimation of moments of the function counting the number of ways to represent integers as a difference of two elements from a $B_2[2]$ set, that

$$F_{2,2}(N) \leq \sqrt{6N} + 1 \sim 2.4495\ldots\sqrt{N},$$

which improved on (15) giving $2.6361\ldots$. Inspired by Cilleruelo's proof, the author then introduced a more elaborate approach [31] using Erdős-Túran-like estimates and the complete Cauchy identity. This gave

$$F_{2,2}(N) \lesssim 2.3636N,$$

an estimate which was refined another time by L. Habsieger and the author [15] to

$$F_{2,2}(N) \lesssim 2.3218\ldots\sqrt{N}.$$

This result was obtained with purely combinatorial tools. We explicitly solve (asymptotically) an integer program corresponding to a limit case.

Note that both Cilleruelo's and the author's results can be generalized to $B_2[g]$ sets. Cilleruelo's theorem reads as

$$F_{2,g}(N) \leq \sqrt{4g-2}\sqrt{N},$$

while Plagne's [31] (which contains a numerical part in connection with an optimization part) furnishes

$$F_{2,g}(N) \leq \begin{cases} 3.04\sqrt{N} \text{ if } g = 3, \\ 3.61\sqrt{N} \text{ if } g = 4, \\ 4.11\sqrt{N} \text{ if } g = 5, \\ 4.53\sqrt{N} \text{ if } g = 6. \end{cases}$$

When $g \geq 7$, there is no more improvement by this method compared to (15): we obtain 4.935 while (15) gives 4.931. Maybe the method of [15] can also be applied in these cases.

Green's technique already mentioned [14] applies also in this case: it improves on all the estimates of this section. Green's results are still the best known upper bounds. It yields

$$F_{2,g}(N)N^{-1/2} \lesssim \min\left(\sqrt{\frac{7g}{2}-\frac{7}{4}}, \sqrt{\frac{17g}{5}}\right), \tag{16}$$

the first value being the minimal one if $g \leq 18$. In any case, inequality (16) improves the result of Cilleruelo, Ruzsa and Trujillo (in fact the second minoration reuses some part of the argument from [9]).

3.3. **The case $h > 2$.** In this case, until very recently nothing else than (3) was known. It has been improved in [9] up to

$$F_{h,g}(N)N^{-1/h} \lesssim \left(\frac{ghh!}{1+\cos^h(\pi/h)}\right)^{1/h}. \tag{17}$$

This relation generalizes (15). Specializing this result (in view of the forthcoming table), we get, for $h = 3$,

$$F_{3,g}(N)N^{-1/3} \lesssim (16g)^{1/3}$$

and for $h = 4$,

$$F_{4,g}(N)N^{-1/4} \lesssim (384g/5)^{1/4}.$$

If this was proved by Fourier arguments, reasonings from probability theory can also be applied. The idea of using such kind of reasoning is not new in the context of combinatorial number theory and even in Sidon-type problems (see Chapter 3 of [17]). But, as far as I know, Alon (see [23]) was the first to introduce a probabilistic point of view for the very precise problem we investigate here (large finite $B_h[g]$ sets). In a presently unpublished preprint [8], Cilleruelo and Jiménez-Urroz present another similar probabilistic argument. Let us sketch it: let X_i $(1 \leq i \leq h)$ be independent random variables uniformly distributed in $\mathcal{A}$, the $B_h[g]$ set to study. Define $Y = X_1 + \cdots + X_h$. General facts show that

$$E((Y-\bar{Y})^2) \leq h\frac{(N-1)^2}{4}.$$

Now, using the fact that Y cannot be too concentrated on its mean-value $\bar{Y}$ in view of the $B_h[g]$ property, what can be obtained is a lower bound for $E((Y-\bar{Y})^2)$ in term of g, h and $|\mathcal{A}|$. Comparing these two bounds, we get

$$F_{h,g}(N)N^{-1/h} \lesssim (g\sqrt{3h}h!)^{1/h}, \tag{18}$$

a new upper bound when $h \geq 7$. A precise analysis of the proof shows that this can be further improved, for example by taking into account information on the repartition of the elements of $\mathcal{A}$ in $\{1, \ldots, N\}$. It would be interesting to optimize the method in order to know by which constant we could replace $\sqrt{3}$ in formula (18). The bound $(g\sqrt{h}h!)^{1/h}$ should be true since it would follow from the equirepartition of elements in the asymptotically maximal $B_h[g]$ sets.

3.4. **Future.** We did not present any specific problem depending on the values of g and h. To my understanding, the real problem is the following.

Problem 5. *Taking advantage of the different approaches presented in the preceding sections (purely combinatorial, probabilistic, analytical, etc.), improve on the upper bounds for $B_h[g]$ sets.*

4. Conclusion and perspectives

To be slightly more explicit, we begin this conclusion with the best known (truncated) asymptotic lower and upper bounds for the quantity $F_{h,g}(N)N^{-1/h}$ for the smallest values of h and g. References to the paper where it is obtained first are given, [x] is for a reference to the section 2.2 of the present paper.

		h		
		2	3	4
g	1	[36] 1 [12]	[4] 1 - 1.5183 [14]	[4] 1 - 1.6266 [14]
	2	[15] 1.5118 - 2.2913 [14]	[4] 1 -3.1748 [9]	[4] 1 - 3.5204 [9]
	3	[9] 1.7888 - 2.9581 [14]	[8] 1.5874 - 3.6342 [9]	[4] 1 - 3.8960 [9]
	4	[x] 2.1552 - 3.5 [14]	[27] 1.5874 - 4 [9]	[4] 1 - 4.1865 [9]
	5	[9] 2.3333 - 3.9687 [14]	[27] 1.5874 - 4.3088 [9]	[4] 1 - 4.4267 [9]
	6	[x] 2.6832 - 4.3875 [14]	[x] 2.0745 - 4.5788 [9]	[8] 1.6817 - 4.6331 [9]

Naturally what remains to be solved is an order of magnitude more difficult than what has already been solved. Anyway, papers on the subject (or related to) have flourished lately and it is likely that new progress should be made soon.

We already wrote down some particular problems in the preceding sections. Of course, the main unsolved question is:

Problem 6. *Does*

$$c_{h,g} = \lim_{N \to +\infty} \frac{F_{h,g}(N)}{N^{1/h}} \tag{19}$$

exist ?

The temptation to answer yes to this question is huge, not only because of the case of Sidon sets but also in view of some "numerical evidences".

A parallel question is

Problem 7. *Compute explicitly $c_{h,g}$ (or if you don't think that this value exists, compute optimal upper and lower bounds).*

It is of course a very difficult task. It was my viewpoint to make my utmost to answer this question in the case of $B_2[2]$ because it should be the simplest case. Unfortunately, even in this case, we are still far from the truth ... Anyway, since the situation can be more efficiently handled in special cases, it would be already nice to answer

Problem 8. *Compute explicitly $c_{3,1}$ and $c_{2,2}$.*

It seems to me that these are the natural cases to study: indeed $B_3[1]$ and $B_2[2]$ sets are representative of the two difficulties in Sidon's problem, namely the number of summands and the number of repetitions allowed. As regards $B_2[2]$ sets, in [15] we risked the conjecture that $c_{2,2} = 2$. I would be pleased to see this conjecture proved ... or disproved because it would mean that we are really near from the right value (in view of numerical evidences). It would be interesting to build optimal $B_3[1]$ sets, so we are lead to the easier problem.

Problem 9. *Estimate conjecturally $c_{3,1}$ (is $c_{3,1} = 1$ reasonable?) or at least find an efficient algorithm to compute the largest $B_3[1]$ set in $\{1, \dots, N\}$.*

On the other side of the spectrum of questions, we ask:

Problem 10. *Find an asymptotic expansion of $c_{h,g}$ when g and h are large.*

We close this paper with the following reaffirmation: we strongly believe that advances are to be expected from mixing different approaches together, namely the combinatorial approach, the Fourier technique and the probabilistic viewpoint. Anyway, we think that many efforts are to be done to touch the truth and that time will be needed: solving the general Sidon's problem seems to be even more difficult than it was to enter Sidon's office: following Pinkus [30] and Babai [1], Erdős was once answered after knocking at Sidon's office door the funny "Please come at some other time, and to someone else" ...

References

[1] L. Babai, C. Pomerance, P. Vértesi, *The mathematics of Paul Erdős*, Notices Amer. Math. Soc. **45** (1998), 19 – 31.

[2] L. Babai, V. T. Sós, *Sidon sets in groups and induced subgraphs of Cayley graphs*, European J. Combin. **6** (1985), 101 – 114.

[3] R. C. Bose, *An affine analogue of Singer's theorem*, J. Indian Math. Soc. **6** (1942), 1 – 15.

[4] R. C. Bose, S. Chowla, *Theorems in the additive theory of numbers*, Comment. Math. Helv. **37** (1962/1963), 141 – 147.

[5] S. Chen, *On the size of finite Sidon sequences*, Proc. Amer. Math. Soc. **121** (1994), 353 – 356.

[6] J. Cilleruelo, *An Upper Bound for $B_2[2]$ Sequences*, J. Combin. Theory Ser. A **89** (2000), 141 – 144.

[7] J. Cilleruelo, *New Upper Bounds for finite B_h Sequences*, preprint (2000).

[8] J. Cilleruelo, J. Jiménez-Urroz, *$B_h[g]$ Sequences*, preprint (2000).

[9] J. Cilleruelo, I. Ruzsa, C. Trujillo, *Upper and lower bounds for finite $B_h[g]$ sequences, $g > 1$*, J. Number Theory, in press.

[10] P. Erdős, *On a problem of Sidon in additive number theory and some related problems, Addendum*, J. London Math. Soc. **19** (1944), 208.

[11] P. Erdős, R. Freud, *On Sidon sequences and related problems*, Mat. Lapok **2** (1991), 1 – 44.

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

[13] S. W. Graham, *B_h sequences*, in Analytic number theory, **1** (Allerton Park, IL, 1995), 431 – 449.

[14] B. J. Green, *The number of squares and $B_h[g]$ sets*, to appear in Acta Arith.

[15] L. Habsieger, A. Plagne, *Ensembles $B_2[2]$: l'étau se resserre*, submitted (2000).

[16] D. Hajela, *Some remarks on $B_h[g]$ sets*, J. Number Theory **29** (1988), 311 – 323.

[17] H. Halberstam, K. F. Roth, *Sequences*, Oxford University Press, 1966.

[18] M. Hall, *Cyclic projective planes*, Duke Math. J. **14** (1947), 1079 – 1090.

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

[20] X.-D. Jia, *$B_h[g]$ sequences with large upper density*, J. Number Theory **56** (1996), 298 – 308.

[21] D. Jungnickel, A. Pott, *Difference sets: an introduction*, in Difference sets, sequences and their correlation properties (Bad Windsheim, 1998), 259–295, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 542, Kluwer Acad. Publ., Dordrecht, 1999.

[22] M. Kolountzakis, *The density of $B_h[g]$ Sequences and the Minimum of Dense Cosine Sums*, J. Number Theory **56** (1996), 4 – 11.

[23] M. Kolountzakis, *Problems in the Additive Number Theory of General Sets, I. Sets with distinct sums*, unpublished manuscript (1996).

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

[25] B. Lindström, *An inequality for B_2 sequences*, J. Combin. Theory **6** (1969), 211 – 212.

[26] B. Lindström, *A remark on B_4 sequences*, J. Combin. Theory **7** (1969), 276 – 277.

[27] B. Lindström, *$B_h[g]$-sequences from B_h sequences*, Proc. Amer. Math. Soc. **128** (1999), 657 – 659.

[28] J. M. López, K. A. Ross, *Sidon sets*, Lecture Notes in Pure and Applied Mathematics **13**, Dekker, New York, 1975.

[29] A. Mian, S. Chowla, *On the B_2-sequences of Sidon*, Proc. Nat. Acad. Sci. India, Sect. A **14** (1944), 3 – 4.

[30] A. Pinkus, *Paul Erdős*, manuscript, available on-line at the address `http://www.math.technion.ac.il/newmath/erdos`.

[31] A. Plagne, *A new upper bound for the cardinality of $B_2[2]$ sets*, J. Combin. Theory Ser. A. **93** (2001), 380 – 386.

[32] I. Z. Ruzsa, *An infinite Sidon sequence*, J. Number Theory **68** (1998), 63 – 71.

[33] I. Z. Ruzsa, *Solving a linear equation in a set of integers I*, Acta Arith. **LXV.3** (1993), 259 – 282.

[34] A Sárközy, V. T. Sós, *On additive representation functions*, The mathematics of Paul Erdős, I, 129–150, Algorithms Combin. **13**, Springer, Berlin, 1997.

[35] S. Sidon, *Ein Satz über trigonometrische Polynome und seine Anwendung in der Theorie des Fourier-Reihen*, Math. Ann. **106** (1932), 536 – 539.

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

Alain Plagne
LIX
École polytechnique
91128 Palaiseau Cedex
France
E-mail address: `plagne@lix.polytechnique.fr`