SUMS AND DIFFERENCES OF FINITE SETS

KATALIN GYARMATI, FRANÇOIS HENNECART, AND IMRE Z. RUZSA

To Jean-Marc Deshouillers, for his 60th birthday

Abstract. In a given abelian group, let A and B be two finite subsets satisfying the small sumset condition $|A+B| \leq K|A|$. We consider the problem of estimating how large $|A-B|$ can be in terms of $|A|$ and K and the one of estimating the ratio $|X-B|/|X|$ when X runs over all the non-empty subsets of A.

1. Introduction and statement of the results

Let A and B be two non-empty and finite subsets of an abelian group G. The cardinality of any finite set X is written $|X|$. As usual, we denote by $A+B$ (resp. $A-B$) the set of all sums $a+b$ (resp. differences $a-b$) where $a \in A$ and $b \in B$. The set of all sums of h elements of B is denoted by hB. In the last fifteen years, several papers were concerning with the problem of comparing the relative sizes of $A+B$ and $A-B$. We clearly have $\max(|A|,|B|) \leq |A \pm B| \leq |A|\,|B|$. The upper bound is achieved when A and B are generic sets, that is when the only solutions of $a+b=a'+b'$, $a,a' \in A$, $b,b' \in B$ are the trivial solutions $(a,b)=(a',b')$. This shows that there is no non-trivial solution for $a-b'=a'-b$, $a,a' \in A$, $b,b' \in B$, thus we also have $|A-B| = |A|\,|B|$. If $|A+B|=|A|$, then $A+B-B=A$, which implies $|A-B|=|A|$. In this paper we consider the question of comparing the size of $A-B$ with that of $A+B$ when $|A+B| \leq K|A|$.

For multiple addition or difference, sharp results have been obtained thanks to a very efficient theorem of Plünnecke. According to [4], this result known as Plünnecke inequalities, can be stated as follows:

(i) *Assume that $|A+B| \leq K|A|$. Then for any positive integer h, there exists a non-empty subset X of A such that*

$$|X+hB| \leq K^h|X|. \tag{1}$$

Date: May 30, 2008.

1991 *Mathematics Subject Classification.* 11B75.

Key words and phrases. Sumset, Difference set, Plünnecke inequality.

The research of the first-named author is partially supported by Hungarian National Foundation for Scientific Research (OTKA), Grants No. T043631, T043623 and T049693.

The research of the third-named author is supported by Hungarian National Foundation for Scientific Research (OTKA), Grants No. T042750, T043623 and T061908.

(ii) *Assume that for a positive integer j one has $|A + jB| \le K|A|$. Then for any integer $h \ge j$, there exists a non-empty subset X of A such that*

$$|X + hB| \le K^{h/j}|X|. \tag{2}$$

(iii) *Assume that $|A + B| \le K|A|$. Then for any nonnegative integers h, j, one has*

$$|hB - jB| \le K^{h+j}|A|.$$

Assertion (i) is a particular case of (ii) and assertion (iii) is obtained by using (ii) and the inequality (cf. [4])

$$|X - Y| \le \frac{|X + Z||Y + Z|}{|Z|}, \tag{3}$$

which is valid for any finite sets X, Y, Z. It is quite clear that in general the set X in (i) and (ii) of Plünnecke inequalities cannot be reduced to a singleton (just think $A = B$ being a large finite arithmetic progression). On the other hand, it is worth mentioning that in general one cannot take $X = A$ (see [6] for more details on this question).

Letting $j = 0$ and $h = 2$ in assertion (iii) of Plünnecke inequality, we obtain $|2B| \le |A + B|^2/|A|$. Thus we have

$$|A - B| \le \frac{|A + B||2B|}{|B|} \le \frac{|A + B|^3}{|A||B|} = \left(\frac{|A + B|^2}{|A||B|}\right)|A + B|, \tag{4}$$

by using inequality (3). When $|A|$, $|B|$ and $|A + B|$ are of comparable size, this inequality shows that $|A - B|$ has also a bounded ratio with $|A|$. If we only assume that $|A + B| \le K|A|$, it is not true that $|A - B|/|A|$ is bounded by some constant depending on K, except in the special case $K = 1$. Indeed, the third-named author proved in [6] the following result: *There exists a real number $\theta > 1$ such that for any $K > 1$ and arbitrarily large integers n, there are two sets of integers A and B satisfying*

$$|A| = n, \quad |A + B| \le K|A| \quad \text{and} \quad |A - B| \ge c(K)|A + B|^\theta, \tag{5}$$

where $c(K) > 0$.

The discussion above shows that the only way to extend this statement to $K = 1$ is to let $c(1) = 0$.

As shown in [6], the choice $\theta = 2 - \frac{\log 6}{\log 7} = 1.0792\ldots$ is admissible in (5). The proof is based on a elementary construction which uses the fact that the set $U = \{0, 1, 3\}$ satisfies $|U + U| = 6$ and $|U - U| = 7$. In this connection and for future references we notice that (3) yields

$$|U - U| \le |U + U|^{4/3}. \tag{6}$$

In [2], it is shown that for any $\lambda < \frac{\log(1+\sqrt{2})}{\log 2} = 1.2715\ldots$, there exist sets A of nonnegative integers such that $|A - A| \asymp |A + A|^\lambda$, but A does not fulfill the condition $|A + A| \asymp |A|$ any more. Nevertheless these sets allow us to show that the exponent θ in (5) can be slightly improved as regards to the original result:

Theorem 1. *Let $K > 1$ be a real number. There exist a real number $\theta_0 > 1.14465$ and two sets of integers A and B with $|A|$ arbitrarily large such that*

$$|A+B| \le K|A| \quad \text{and} \quad |A-B| \ge \left(\frac{2(K-1)}{3K}\right)^{5/4} |A+B|^{\theta_0}. \tag{7}$$

Using similar ideas, one can show that there exists a positive real number $c(K)$ such that for any positive integer n, there exists two sets of integers A and B for which (5) holds with $\theta = \theta_0$.

The easy bound $|2B| \le |B|^2$ and (4) imply $|A-B| \le |A+B||2B|^{1/2}$. Since $|3B|^{1/3} \le |2B|^{1/2}$ (see [6, Theorem 7.2] and also [7]), the following result provides a strenghtened estimate.

Theorem 2. *Let A and B two finite sets in an abelian group. Then*

$$|A-B| \le |A+B|\,|3B|^{1/3}. \tag{8}$$

In [7], the third-named author suggested that perhaps, the sequence $(|hB|^{1/h})_{h\ge 1}$ is non-increasing. A natural problem is to find for which integers h we have

$$|A-B| \le |A+B||hB|^{1/h} \tag{9}$$

for any sets A and B. Assume that this bound holds for some $h \ge 1$. By Plünnecke inequality, we have $|hB| \le K^{h-1}|A+B|$, where $K = |A+B|/|A|$. Therefore $|A-B| \le K^{1-1/h}|A+B|^{1+1/h}$. This contradicts Theorem 1 for $h \ge 7$ (see also the remark at the end of Section 2).

Using the trivial fact that $|A||B| \ge |A-B|$, the bound $|A-B| \le |A+B|^{3/2}$ follows from (4). This estimate can be strengthened if we further assume that $|A+B| \le K|A|$:

Corollary 3. *Let A and B be two finite sets such that $|A+B| \le K|A|$. Then*

$$|A-B| \le K^{2/3}|A+B|^{4/3}.$$

Indeed, as $|3B| \le |A+B|^3/|A|^2$ by Plünnecke inequality, Theorem 2 gives

$$|A-B| \le \frac{|A+B|^2}{|A|^{2/3}} \le K^{2/3}|A+B|^{4/3}.$$

From Corollary 3 we deduce that the value of θ in (5) and that of θ_0 in (7) cannot be larger than 4/3.

We now consider the following related question: under the same assumption $|A+B| \le K|A|$, how large can be $|X-B|/|X|$ where X runs over all the subsets of A? Using Plünnecke inequality (1), it is possible to obtain the following upper bound for this ratio:

Theorem 4. *Let A and B be non-empty and finite subset of some abelian group such that $|A+B| \le K|A|$. Then there exists some non-empty subset X of A such that*

$$\frac{|X-B|}{|X|} \le K \exp\left(2\sqrt{(\log K)(\log |A|)}\right). \tag{10}$$

We observed above that $|A-B|/|A|$ can be very large even in the case where $|A+B|/|A|$ is bounded. The following result shows that this fact is in some sense uniform (see [6]): *There exist two sets A and B with $|A|$ arbitrarily large and $|A+B|\le 3|A|$ such that for any $X\subset A$, one has* $|X-B|\ge \frac{1}{3}(\log|A|)|X|$. By a modification of the argument, this result may be improved in the following way:

Theorem 5. *Let $K>1$ and τ such that $0<\tau<1-1/K$, and define*

$$f(\tau)=(-\tau\log\tau-(1-\tau)\log(1-\tau)).$$

Then for any $c<\sqrt{\frac{2}{3}f(\tau)}$ there exist two sets A and B with $|A|$ arbitrarily large and $|A+B|\le K|A|$ such that for any non-empty subset X of A, one has

$$\frac{|X-B|}{|X|}\ge\exp\left(c\sqrt{(\log((1-\tau)K))(\log|A|)(\log\log|A|)^{-1}}\right).$$

As an immediate consequence, we obtain for K not too close to 1:

Corollary 6. *Let $K>2$. Then for any $c<\frac{\sqrt{2}\log 2}{\sqrt{3}}$, there exist two sets A and B with $|A|$ arbitrarily large and $|A+B|\le K|A|$ such that for any non-empty subset X of A, one has*

$$\frac{|X-B|}{|X|}\ge\exp\left(c\sqrt{(\log(K/2))(\log|A|)(\log\log|A|)^{-1}}\right).$$

This uniform lower bound for $|X-B|/|X|$ can be compared to the upper bound (10) obtained in Theorem 4.

Acknowledgement. We are grateful to the referee for his valuable comments and suggestions helping us to improve on the submitted version of the present paper.

2. Sumset and difference set

Proof of Theorem 1. The result will follow from

Lemma. *Let $K>1$ be a real number and let U be a finite, non-empty set of nonnegative integers containing 0. Set $s=|2U|$, $d=|U-U|$, $q=2\max U+1$ and $\theta=1+\log(d/s)/\log q$. If $d<q$, then there exist pairs (A,B) of finite, non-empty integer sets with $|B|$ arbitrarily large such that $|A+B|\le K|A|$ and*

$$|A-B|\ge(2(K-1)/3K)^{5/4}|A+B|^{\theta}. \tag{11}$$

Proof. We fix k any arbitrary large integer. Set

$$B=\left\{\sum_{j=0}^{k-1}u_jq^j\ :\ u_j\in U,\ j=0,\dots,k-1\right\},$$

and

$$A=[1,L]\cup\bigcup_{i=1}^{m}(a_i+B),$$

where the a_i's are positive integers larger than $L+q^k$ and such that $a_i-a_j \notin (B-B)\cup 2B$ unless $i=j$. Since $\max B < q^k$, we have

$$|A| \geq L+1, \quad |A+B| = ms^k + t, \quad |A-B| = md^k + t,$$

where $t := |[1,L]+B| = |[1,L]-B|$. Since $B \subset [0, \frac{q^k}{2}]$, we note that $L \leq t \leq L + \frac{q^k}{2}$. We choose

$$L = \left\lfloor \frac{3q^k}{2(K-1)} \right\rfloor.$$

Letting $m = \lfloor \left(\frac{q}{s}\right)^k \rfloor$, we obtain $|A+B| \leq q^k + t \leq \frac{3}{2}q^k + L \leq \frac{3Kq^k}{2(K-1)} \leq K(L+1) \leq K|A|$ and $|A-B| \geq (\frac{qd}{s})^k - d^k + t \geq (\frac{qd}{s})^k$ if we assume further that $d < q$ and k is sufficiently large. Consequently

$$|A-B| \geq \left(\frac{qd}{s}\right)^k \geq \left(\frac{2(K-1)|A+B|}{3K}\right)^{1+\frac{\log d - \log s}{\log q}}.$$

By (6), we have $d \leq \max(q, s^{4/3})$, thus

$$1 + \frac{\log d - \log s}{\log q} \leq \frac{5}{4}. \tag{12}$$

We finally get (11). $\square$

Remark. It is worth mentioning that (12) implies that the largest exponent θ that could be eventually obtained by this method is at most equal to $5/4$.

By an exhaustive computational research, we got the set $U = \{0, 1, 3, 6, 13, 17, 21\}$ which satisfies $|U+U| = 26$, $|U-U| = 39$ and $q = 43$, thus the exponent $\theta = 1 + \frac{\log 39 - \log 26}{\log 43} = 1.1078\ldots$ is admissible in (5) with $c(K) = \left(\frac{2(K-1)}{3K}\right)^{5/4}$. This set U provides the optimal value of $\frac{\log d(U) - \log s(U)}{\log q(U)}$ when U runs over all sets of nonnegative integers of cardinality less than or equal to 11.

In order to improve the admissible exponent in (5), we will use some idea from [2]. We denote $\mathbb{N}$ the set of all nonnegative integers. Let

$$V = V(m,L) = \{(x_1, \ldots, x_m) \in \mathbb{N}^m \,:\, x_1 + \cdots + x_m \leq L\}. \tag{13}$$

Then by lemmas 1 and 2 of [2], we get

$$|V| = \binom{m+L}{m}, \; |2V| = \binom{m+2L}{m}, \; |V-V| = \sum_{k=0}^{\min(m,L)} \binom{m}{k}^2 \binom{L+m-k}{m}. \tag{14}$$

Let $\Lambda = (L_j)_{j\geq 0}$ be the sequence defined by

$$L_0 = 1, \quad L_{j+1} = 2LL_j + 1, \; j \geq 0. \tag{15}$$

By projection of V on the set of nonnegative integers $(x_1, \ldots, x_m) \mapsto x_1 + x_2L_1 + x_3L_2 + \cdots + x_mL_{m-1}$, by which the number of sums and the number of differences are preserved, we get a set U verifying $\max U = LL_{m-1}$. Solving the linear recurrence (15), we obtain $L_{m-1} = \frac{(2L)^m-1}{2L-1}$, thus $q(U) = 2\max U + 1 = \frac{(2L)^{m+1}-1}{2L-1}$. The choice $m = 8$, $L = 9$ gives a set U with $|U| = 24310$, $s(U) = 1562275$, $d(U) = 23301307$ and $q(U) = 11668193551$. This yields the exponent

$$\theta = 1 + \frac{\log d(U) - \log s(U)}{\log q(U)} = 1.1165\dots$$

in (5).

We may observe that when projecting V on the set of integers, we only need to select a sequence $\Lambda = (L_j)_{j=0,\dots,m-1}$ such that the number of sums (and hence also the number of differences) are preserved. For this we can argue by induction applying the following greedy algorithm: let $L_0 = 1$, and assume that for some $1 \le j \le m-1$, $L_0 < L_1 < \cdots < L_{j-1}$ have been chosen so that the mapping $p_j : (x_1,\dots,x_j) \mapsto x_1 + x_2L_1 + x_3L_2 + \cdots + x_jL_{j-1}$ preserves the number of sums from $S(j,L) := \{(x_1,\dots,x_j) \in \mathbb{N}^j : x_1 + \cdots + x_j \le L\}$. Put $U(j,L) := p_j(S(j,L))$ and let

$$L_j := \min\{l > LL_{j-1} : l \notin U(j,L) + U(j,L) - U(j,L) - U(j,L-1)\}.$$

Then the projection $p_{j+1} : (x_1,\dots,x_{j+1}) \mapsto x_1 + x_2L_1 + x_3L_2 + \cdots + x_jL_{j-1} + x_{j+1}L_j$ preserves the number of sums from $S(j+1,L)$. Indeed let $x,y,z,t \in S(j+1,L)$ such that

$$p_{j+1}(x) + p_{j+1}(y) = p_{j+1}(z) + p_{j+1}(t). \tag{16}$$

If $x_{j+1} = y_{j+1} = z_{j+1} = t_{j+1} = 0$, then

$$p_j(x_1,\dots,x_j) + p_j(y_1,\dots,y_j) = p_j(z_1,\dots,z_j) + p_j(t_1,\dots,t_j), \tag{17}$$

hence by induction hypothesis $x + y = z + t$. Otherwise, we may assume that $x_{j+1} + y_{j+1} - z_{j+1} - t_{j+1} \ge 0$ and $x_{j+1} \ge 1$. Then $(x_1,\dots,x_j) \in S(j,L-1)$ and by (16), one has $(x_{j+1} + y_{j+1} - z_{j+1} - t_{j+1})L_j = p_j(t_1,\dots,t_j) + p_j(z_1,\dots,z_j) - p_j(y_1,\dots,y_j) - p_j(x_1,\dots,x_j) \in U(j,L) + U(j,L) - U(j,L) - U(j,L-1)$. Since $\max(U(j,L) + U(j,L) - U(j,L) - U(j,L-1)) < 2L_j$ and $L_j \notin U(j,L) + U(j,L) - U(j,L) - U(j,L-1)$, we clearly have $x_{j+1} + y_{j+1} - z_{j+1} - t_{j+1} = 0$, giving (17) again. By the induction hypothesis, we deduce $(x_1,\dots,x_j) + (y_1,\dots,y_j) = (z_1,\dots,z_j) + (t_1,\dots,t_j)$, and finally $x + y = z + t$. For $m = 9$ and $L = 7$, a short program gives the sequence

$$\Lambda = (1, 15, 211, 1590, 14976, 109870, 788046, 5535439, 38772709)$$

yielding by projection a sequence U of integers such that $q(U) = 2\max U + 1 = 542817927$. Since sums and differences are preserved in cardinality, of course by (14) we have $s(U) = \binom{23}{9} = 817190$ and $d(U) = \sum_{k=0}^{6} \binom{9}{k}^2 \binom{16-k}{9} = 12494233$. We thus get $\theta = 1.135596$ as an admissible exponent.

It is still possible to improve it by relaxing the definition of the sequence $\Lambda = (L_j)_{j=0,\dots,m-1}$ by removing the condition $L_j > LL_{j-1}$, $j \ge 1$. We thus obtain a new sequence Λ for which the projection $p_j : (x_1,\dots,x_j) \mapsto x_1 + x_2L_1 + x_3L_2 + \cdots + x_jL_{j-1}$ does not necessary preserve the number of sums nor the number of differences. However only a few number of sums and differences are lost through the projection p_j. This gives for $m = 11$, $L = 7$ and

$$\Lambda = (1, 15, 211, 1590, 14976, 109870, 605315, 3362489, 17767138, 80137194, 408850463)$$

a set U verifying

$$s(U) = 4455634, \quad d(U) = 110205905, \quad q(U) = 2 \max U + 1 = 5723906483.$$

This yields the admissible exponent $\theta = 1.144655$. □

Proof of Theorem 2. The Plünnecke inequality (i) given in the introduction has the disadvantage not to give any information on the size of the subset X of A. However by repeated application of it, it has been shown by the third-named author that an analogue result holds with a large subset X of A (see [7, Theorem 3.3]). In a weaker but more convenient form, it can be stated as follows:

Lemma. *Let K and δ be positive real numbers, h be a positive integer and A, B be finite and non-empty subsets of an abelian group such that $|A+B| \le K|A|$. Then there exists a subset X of A with $|X| \ge (1-\delta)|A|$ such that $|X + hB| \le 2K^h \delta^{1-h}|A|$.*

We now complete the proof of Theorem 2. We use the following notation: $|A| = m$, $|jB| = n_j$, $|B| = n = n_1$, $|A + B| = s$ and $|A - B| = d$. We obviously have

$$d \le mn. \tag{18}$$

We also use several instances of (3). First we put $X = A$, $Y = B$, $Z = B$ to obtain

$$d \le \frac{sn_2}{n}. \tag{19}$$

Next we put $Y = B$, $Z = 2B$ to obtain

$$|X - B| \le |X + 2B| \frac{n_3}{n_2}. \tag{20}$$

We will use this for a large subset X of A for which $X + 2B$ is small and in view of (20) we will then estimate $A - B$ by

$$|A - B| \le |X - B| + |(A \smallsetminus X) - B| \le |X + 2B| \frac{n_3}{n_2} + n(m - |X|).$$

For the set X given in the lemma with $h = 2$, we deduce

$$|A - B| \le \frac{2n_3 s^2}{n_2 \delta m} + \delta nm. \tag{21}$$

Choosing $\delta = \frac{s}{m}\left(\frac{2n_3}{nn_2}\right)^{1/2}$ in this inequality, we find

$$|A - B|^2 \le (2s)^2 \left(\frac{2nn_3}{n_2}\right).$$

Multiplying this inequality with (19) and taking the cube root, we obtain $d \le 2sn_3^{1/3}$, which is the requested inequality apart from the factor 2. We can remove it as follows.

Take our sets A, B and apply the result to the k-fold Cartesian products A^k and B^k. Every quantity is then raised to the k-th power, and by taking k-th root we have our theorem with the factor $2^{1/k}$. By taking the limit we derive the theorem with the factor 1. $\square$

Remark. We saw in the introduction that the bound (9) is not true in general for $h \geq 7$. Let $A = B = V(m, m/2)$ be the set defined in (13) with $L = m/2$. We have by (14) the estimates $\log|2A| = (2\log 2 + o(1))m$, $\log|A - A| = (2\log(1+\sqrt{2}) + o(1))m$ as m tends to infinity (see [2] for more details). Moreover $6A = V(m, 3m)$, thus, by Stirling's formula, we have $|6A| = (4\log 4 - 3\log 3 + o(1))m$ as m tends to infinity. Since $2\log(1+\sqrt{2}) - 2\log 2 > \frac{4\log 4 - 3\log 3}{4}$, we obtain that $|A - A| > |2A||6A|^{1/6}$ for m sufficiently large, disproving the bound (9) for $h = 6$. For $h = 4$ or 5, it is an open question to decide whether or not (9) holds for any sets A and B.

3. How large can $|X - B|$ be for $X \subset A$?

Proof of Theorem 4. For an integer $N \geq 1$ (to be specified later) put

$$\lambda = \min_{1 \leq j \leq N} \frac{|(j+1)B|}{|jB|}.$$

Then by Plünnecke inequality, $\lambda^N |B| \leq |(N+1)B| \leq K^{N+1}|A|$, thus

$$\lambda \leq K^{1+1/N} \left(\frac{|A|}{|B|}\right)^{1/N}.$$

Together with the trivial bound $\lambda \leq |B|$, we get $\lambda \leq K|A|^{1/(N+1)}$. Therefore there exists j, $1 \leq j \leq N$, such that

$$|jB + B| \leq K|A|^{1/(N+1)}|jB|.$$

Inequality (3) yields for any $X \subset A$,

$$|X - B| \leq \frac{|X + jB||(j+1)B|}{|jB|}.$$

By Plünnecke's theorem, there exists a non-empty subset $X \subset A$ such that $|X + jB| \leq K^j|X|$, thus

$$|X - B| \leq K^{j+1}|A|^{1/(N+1)}|X| \leq K^{N+1}|A|^{1/(N+1)}|X|.$$

Taking

$$N = \left\lceil \left(\frac{\log|A|}{\log K}\right)^{1/2} \right\rceil - 1,$$

we finally obtain the bound (10). $\square$

Proof of Theorem 5. Let $d \geq 1$ be an integer. We will construct a pair of sets A and B in $\mathbb{Z}^d$ satisfying the conclusion of Theorem 5. Then by projection on $\mathbb{Z}$, using for instance the mapping $(x_1, \dots, x_d) \mapsto x_1 + qx_2 + \dots + q^{d-1}x_d$ where q is sufficiently large to have the number of sums and that of differences unchanged, we may obtain the same result with A and B being sets of integers.

For a given d-tuple $\underline{x} = (x_1, x_2, \dots, x_d) \in \mathbb{N}^d$, we denote by $\nu(\underline{x})$ the number its non-zero coordinates, and by $\sigma(\underline{x})$ the sum of all its coordinates:

$$\nu(\underline{x}) = \sum_{\substack{1\leq i\leq d\\ x_i\neq 0}} 1, \quad \sigma(\underline{x}) = \sum_{1\leq i\leq d} x_i.$$

Let $(e_i)_{1\leq i\leq d}$ be the canonical basis of $\mathbb{Z}^d$ and $u \in [1, d]$ be an integer. We let

$$A = \{\underline{x} = (x_1, x_2, \dots, x_d) \in \mathbb{N}^d \;:\; \nu(\underline{x}) = J \text{ and } \sigma(\underline{x}) = k\},$$

and

$$B = \{e_{i_1} + e_{i_2} + \dots + e_{i_u} \;:\; 1 \leq i_1 < i_2 < \dots < i_u \leq d\}.$$

The set A is formed with integral points of certains J-dimensional edges of a simplex and the set B by some vertices of an hypercube. The sumset $A + B$ has the same structure than A and its size is controlled by the parameters k and u: large k and small u make $|A + B|$ close to $|A|$. Now each element of $A - B$ having exactly u negative coordinates (all are equal to -1) belongs to a certain $a - B$, for an unique $a \in A$. It follows that choosing the parameter $d - J$ as large as possible, in relation with k and u, will imply a large lower bound for $|X - B|/|X|$, for any $\varnothing \neq X \subset A$.

We have by easy combinatorical considerations

$$|A| = \binom{d}{J}\binom{k-1}{J-1}. \tag{22}$$

Put for $i = 0, 1, \dots, u$

$$C_i = \{\underline{x} = (x_1, x_2, \dots, x_d) \in \mathbb{N}^d \;:\; \nu(\underline{x}) = J + i \text{ and } \sigma(\underline{x}) = k + u\}.$$

Then $A + B \subset \bigcup_{i=0}^{u} C_i$. We also have

$$|C_i| = \binom{d}{J+i}\binom{k+u-1}{J+i-1}.$$

From this and (22) we get

$$\begin{aligned}\frac{|C_i|}{|A|} = &\frac{(d-J)(d-J-1)\dots(d-J+i-1)}{(J+1)(J+2)\dots(J+i)} \cdot \frac{(k+u-1)(k+u-2)\dots(k+u-i)}{(J+i-1)(J+i-2)\dots J} \\ &\cdot \frac{(k+u-i-1)(k+u-i-2)\dots k}{(k-J+u-i)(k-J+u-i-1)\dots(k-J+1)} \\ \leq &\left(\frac{d-J}{J}\right)^i \frac{(k+u)^u}{J^i(k-J)^{u-i}}.\end{aligned}$$

Thus

$$\sum_{i=0}^{u} \frac{|C_i|}{|A|} \leq \left(\frac{k+u}{k-J}\right)^u \sum_{i=0}^{u} \left(\frac{(d-J)(k-J)}{J^2}\right)^i.$$

If we assume

$$\frac{(d-J)(k-J)}{J^2} \leq \tau, \tag{23}$$

we get

$$\frac{|A+B|}{|A|} \leq \sum_{i=0}^{u} \frac{|C_i|}{|A|} \leq (1-\tau)^{-1} \left(\frac{k+u}{k-J}\right)^u. \tag{24}$$

For each $\underline{x} \in A$, there are $(d-J)$ zero coordinates x_i, thus there are at least $\binom{d-J}{u}$ elements in $\underline{x} - B$ which are uniquely determined by $\underline{x}$ in $A - B$. This gives for any $X \subset A$

$$|X - B| \geq \binom{d-J}{u} |X|.$$

We now come to the choice of the parameters. Let $\varepsilon > 0$ such that $(1-\tau)K^{1-\varepsilon} > 1$. We introduce $\theta = \left(\frac{\log((1-\tau)K^{1-\varepsilon})}{J}\right)^{1/2}$, $\lambda = \frac{\tau}{\theta}$ and put

$$u = \lfloor \tau\theta J \rfloor, \quad d = \lfloor (1+\theta)J \rfloor, \quad k = \lfloor (1+\lambda)J \rfloor.$$

Condition (23) is clearly fulfilled thus (24) holds. A short calculation yields

$$(1-\tau)^{-1} \left(\frac{k+u}{k-J}\right)^u \leq (1-\tau)^{-1}(1-\tau)K^{1-\varepsilon}(1+o(1)) \leq K$$

as J tend to infinity, thus $|A+B| \leq K|A|$ can be achieved by taking J large enough.

Stirling's formula gives

$$\binom{d-J}{u} = \binom{\lfloor \theta J \rfloor}{\lfloor \tau\theta J \rfloor} \geq \exp\left((f(\tau)+o(1))\theta J\right),$$

as J tends to infinity. Thus we have

$$\frac{|X-B|}{|X|} \geq \exp\left((f(\tau)+o(1))\sqrt{J\log((1-\tau)K^{1-\varepsilon})}\right).$$

By (22), we obtain $|A| \leq J^{\frac{3}{2}J(1+o(1))}$ as J tends to infinity, hence

$$\log|A| \leq \left(\frac{3}{2} + o(1)\right) J \log J, \tag{25}$$

giving $J \geq \frac{2+o(1)}{3} \frac{\log|A|}{\log\log|A|}$. Theorem 5 follows easily by choosing $\varepsilon > 0$ sufficiently small so that $(1-\varepsilon)^{1/2} f(\tau)\sqrt{\frac{2}{3}} > c$ and then by taking J large enough. $\square$

References

[1] Freiman, G. A., Pigarev W. P.: The relation between the invariants R and T (Russian), Kalinin. Gos. Univ. Moscow (1973), 172–174.
[2] Hennecart, F., Robert, G., Yudin, A.: On the number of sums and differences. In *Structure theory of set addition*. Astérisque No. 258 (1999), 173–178.
[3] Ruzsa, I. Z.: On the cardinality of $A+A$ and $A-A$. In *Coll. Math. Soc. Bolyai* **18**, Combinatorics (Keszthely 1976), Akadémiai Kiadó (Budapest 1979), 933–938.
[4] Ruzsa, I. Z.: An application of graph theory to additive number theory. *Scientia* **3** (1989), 97–109.
[5] Ruzsa, I. Z.: On the number of sums and differences. *Acta Math. Hungar.* **59** (1992), 439–447.
[6] Ruzsa, I. Z.: Sums of finite sets. In *Number theory* (New York, 1991–1995), 281–293, Springer, New York, 1996.
[7] Ruzsa, I. Z.: Cardinality questions about sumsets. to appear.

K. Gyarmati, Alfréd Rényi Institute of Mathematics, Budapest, Pf. 127, H-1364 Hungary
E-mail address: gykati@cs.elte.hu

F. Hennecart, LaMUSE, 23 rue Michelon, 42023 Saint-Etienne cedex 2, France
E-mail address: francois.hennecart@univ-st-etienne.fr

I. Z. Ruzsa, Alfréd Rényi Institute of Mathematics, Budapest, Pf. 127, H-1364 Hungary
E-mail address: ruzsa@renyi.hu