Additive combinatorics and geometry of numbers

Imre Z. Ruzsa*

Abstract. We meditate on the following questions. What are the best analogs of measure and dimension for discrete sets? How should a discrete analogue of the Brunn–Minkowski inequality look like? And back to the continuous case, are we happy with the usual concepts of measure and dimension for studying the addition of sets?

Mathematics Subject Classification (2000). Primary 11B75; Secondary 05-99, 11H06, 52C07.

Keywords. Sumsets, additive combinatorics, lattice points, volume.

1. Introduction

"Additive combinatorics" is a name coined by (I think) Tao and Van for the title of their book in preparation to denote the study of additive properties of general sets – mainly of integers, but also in other structures. Works on this topics are generally classified as additive or combinatorial number theory.

The first result that connects additive properties to geometrical position is perhaps the following theorem of Freiman.

Theorem 1.1 (Freiman [3], Lemma 1.14). *Let $A \subset \mathbb{R}^d$ be a finite set, $|A| = m$. Assume that A is proper d-dimensional, that is, it is not contained in any affine hyperplane. Then*

$$|A + A| \geq (d + 1)m - \frac{d(d + 1)}{2}.$$

This theorem is exact, equality can occur, namely it holds when A is a "long simplex", a set of the form

$$L_{dm} = \{0, e_1, 2e_1, \dots, (m - d)e_1, e_2, e_3, \dots, e_d\}. \tag{1.1}$$

In particular, if no assumption is made on the dimension, then the minimal possible cardinality of the sumset is $2m - 1$, with equality for arithmetic progressions.

This result can be extended to sums of different sets. This extension is problematic from the beginning, namely the assumption "d-dimensional" can be interpreted in

*Supported by Hungarian National Foundation for Scientific Research (OTKA), Grants No. T 38396, T 43623, T 42750.

Proceedings of the International Congress
of Mathematicians, Madrid, Spain, 2006
© 2006 European Mathematical Society

different ways. We can stipulate that both sets be d-dimensional, or only one, or, in the weakest form, make this assumption on the sumset only.

An immediate extension of Freiman's above result goes as follows.

Theorem 1.2 ([11], Corollary 1.1). *If $A, B \subset \mathbb{R}^d$, $|A| \le |B|$ and $\dim(A+B) = d$, then we have*

$$|A+B| \ge |B| + d|A| - \frac{d(d+1)}{2}.$$

We can compare these results to the continuous case. Let A, B be Borel sets in $\mathbb{R}^d$; μ will denote the Lebesgue measure. The celebrated Brunn–Minkowski inequality asserts that

$$\mu(A+B)^{1/d} \ge \mu(A)^{1/d} + \mu(B)^{1/d}, \tag{1.2}$$

and here equality holds if A and B are homothetic convex sets, and under mild and natural assumptions this is the only case of equality. It can also be observed that the case $A = B$ is completely obvious here: we have

$$\mu(A+A) \ge \mu(2 \cdot A) = 2^d \mu(A).$$

Also the constant 2^d is much larger than the constant $d+1$ in Theorem 1.1. This is necessary, as there are examples of equality, however, one feels that this is an exceptional phenomenon and better estimations should hold for "typical" sets. A further difference is the asymmetrical nature of the discrete result and the symmetry of the continuous one. Finally, when $|A|$ is fixed, Theorem 1.2 gives a linear increment, while (1.2) yields

$$\mu(A+B) \ge \mu(B) + d\mu(A)^{1/d}\mu(B)^{1-1/d}.$$

In the next section we tell what can be said if we use cardinality as the discrete analog of measure, and prescribe only the dimension of the sets. Later we try to find other spatial properties that may be used to study sumsets.

We meditate on the following questions (without being able to even conjecture a definitive answer). What are the best analogs of measure and dimension for discrete sets? How should a discrete analogue of the Brunn–Minkowski inequality look like? The partial answers also suggest questions in the continuous case. Should we be satisfied with the usual concepts of measure and dimension for studying the addition of sets?

Most of the paper is a survey, however, there are some new results in Sections 4 and 6.

We end the introduction by fixing some notations, which were tacitly used above.

For two sets A, B (in any structure with an operation called addition) by their *sum* we mean the set

$$A + B = \{a + b : a \in A,\ b \in B\}.$$

We use $A - B$ similarly. For repeated addition we write

$$kA = A + \cdots + A \quad (k \text{ times}),$$

in contrast to

$$k \cdot A = \{ka : k \in A\}.$$

Mostly our sets will be in an Euclidean space $\mathbb{R}^d$, and $e_1, \dots, e_d$ will be the system of unit vectors. We define initially the *dimension* $\dim A$ of a set $A \subset \mathbb{R}^d$ as the dimension of the smallest affine hyperplane containing A. (This definition will be modified in Section 3).

2. Results using cardinality and dimension

We consider finite sets in an Euclidean space $\mathbb{R}^d$.

Put

$$\begin{aligned} F_d(m,n) &= \min\{|A+B| : |A| = m, |B| = n,\ \dim(A+B) = d\},\\ F'_d(m,n) &= \min\{|A+B| : |A| = m, |B| = n,\ \dim B = d\},\\ F''_d(m,n) &= \min\{|A+B| : |A| = m, |B| = n,\ \dim A = \dim B = d\}. \end{aligned}$$

F_d is defined for $m+n \ge d+2$, F'_d for $n \ge d+1$ and F''_d for $m \ge d+1$, $n \ge d+1$. F_d and F''_d are obviously symmetric, while F'_d may not be (and, in fact, we will see that for certain values of m, n it is not), and they are connected by the obvious inequalities

$$F_d(m,n) \le F'_d(m,n) \le F''_d(m,n).$$

I determined the behaviour of F_d and of F'_d for $m \le n$. The more difficult problem of describing F''_d and F'_d for $m > n$ was solved by Gardner and Gronchi [4]; we shall quote their results later.

To describe F_d define another function G_d as follows:

$$G_d(m,n) = n + \sum_{j=1}^{m-1} \min(d, n-j), \quad n \ge m \ge 1$$

and for $m > n$ extend it symmetrically, putting $G_d(m,n) = G_d(n,m)$. In other words, if $n - m \ge d$, then we have

$$G_d(m,n) = n + d(m-1).$$

If $0 \le t = n - m < d$, then for $n > d$ we have

$$G_d(m,n) = n + d(m-1) - \frac{(d-t)(d-t-1)}{2} = n(d+1) - \frac{d(d+1)}{2} - \frac{t(t+1)}{2},$$

and for $n \le d$

$$G_d(m,n) = n + \frac{(m-1)(2n-m)}{2}.$$

With this notation we have the following result.

Theorem 2.1 ([11], Theorem 1). *For all positive integers m, n and d satisfying $m+n \geq d+2$ we have*

$$F_d(m,n) \geq G_d(m,n).$$

Theorem 1.2 is an immediate consequence.

Theorem 2.1 is typically exact; the next theorem summarizes the cases when we have examples of equality.

Theorem 2.2 ([11], Theorem 2). *Assume $1 \leq m \leq n$. We have*

$$F_d(m,n) = F'_d(m,n) = G_d(m,n)$$

unless either $n < d+1$ or $m \leq n-m \leq d$ (in this case $n \leq 2d$).

The construction goes as follows.

Assume $1 \leq m \leq n$, $n \geq d+1$. Let B be a long simplex, $B = L_{dn}$ as defined in (1.1).

If $n-m \geq d$, we put

$$A = \{0e_1, 1e_1, \dots, (m-1)e_1\}.$$

This set satisfies $|A| = m$. The set $A+B$ consists of the vectors ie_1, $0 \leq i \leq n+m-d-1$ and the vectors $ie_1 + e_j$, $0 \leq i \leq m-1$, $2 \leq j \leq d$, consequently

$$|A+B| = n + d(m-1) = G_d(m,n).$$

If $n-m = t < d$, write $t = d-k$ and assume $k \leq m$. Now A is defined by

$$A = \{0e_1, 1e_1, \dots, (m-k)e_1\} \cup \{e_2, \dots, e_k\}.$$

This set satisfies $|A| = m$. The set $A+B$ consists of the vectors ie_1, $0 \leq i \leq 2(n-d)$, the vectors $ie_1 + e_j$, $0 \leq i \leq n-d$, $2 \leq j \leq d$, finally $e_i + e_j$, $2 \leq i, j \leq k$, hence

$$\begin{aligned} |A+B| &= 2(n-d) + 1 + (d-1)(n-d+1) + \frac{k(k-1)}{2} \\ &= n(d+1) - \frac{d(d+1)}{2} - \frac{t(t+1)}{2} = G_d(m,n). \end{aligned}$$

These constructions cover all pairs m, n except those listed in Theorem 2.2. Observe that A is also a long simplex of lower dimension. For a few small values the exact bounds are yet to be determined.

We now describe Gardner and Gronchi's [4] bound for $F'_d(m,n)$. Informally their main result (Theorem 5.1) asserts that the $|A+B|$ is minimalized when $B = L_{dn}$, a long simplex, and A is as near to the set of points inside a homothetic simplex as possible. More exactly the define (for a fixed value of n) the weight of a point $x = (x_1, \dots, x_d)$ as

$$w(x) = \frac{x_1}{n-d} + x_2 + \cdots + x_d.$$

This defines an ordering by writing $x < y$ if either $w(x) < w(y)$ or $w(x) = w(y)$ and for some j we have $x_j > y_j$ and $x_i = y_i$ for $i < j$.

Let D_{dmn} be the collection of the first m vectors with nonnegative integer coordinates in this ordering. We have $D_{dnn} = L_{dn} = B$, and, more generally, $D_{dmn} = rB$ for any integer m such that

$$m = |rB| = (n-d)\binom{r+d-1}{d} + \binom{r+d-1}{d-1}.$$

For such values of m we also have

$$|A+B| = |(r+1)B| = (n-d)\binom{r+d}{d} + \binom{r+d}{d-1}.$$

With this notation their result sounds as follows.

Theorem 2.3 (Gardner and Gronchi [4], Theorem 5.1). *If $A, B \subset \mathbb{R}^d$, $|A| = m$, $|B| = n$ and* $\dim B = d$*, then we have*

$$|A+B| \geq |D_{dmn} + L_{dn}|.$$

For $m < n$ this reproves Theorem 2.2. For $m \geq n$ the extremal set D_{dmn} is also d-dimensional, thus this result also gives the value of F''_d.

Corollary 2.4. *For $m \geq n > d$ we have*

$$F''_d(m,n) = F'_d(m,n) = |D_{dmn} + L_{dn}|.$$

A formula for the value of this function is given in [4], Section 6. We quote some interesting consequences.

Theorem 2.5 (Gardner and Gronchi [4], Theorem 6.5). *If $A, B \subset \mathbb{R}^d$, $|A| = m \geq |B| = n$ and* $\dim B = d$*, then we have*

$$|A+B| \geq m + (d-1)n + (n-d)^{1-1/d}(m-d)^{1/d} - \frac{d(d-1)}{2}.$$

Theorem 2.6 (Gardner and Gronchi [4], Theorem 6.6). *If $A, B \subset \mathbb{R}^d$, $|A| = m$, $|B| = n$ and* $\dim B = d$*, then we have*

$$|A+B|^{1/n} \geq m^{1/d} + \left(\frac{n-d}{d!}\right)^{1/d}.$$

This result is as close to the Brunn–Minkowski inequality as we can get by using only the cardinality of the summands.

3. The impact function and the hull volume

While we will focus our attention to sets in Euclidean spaces, some definitions and results can be formulated more clearly in a more general setting. So let now G be a commutative group. For a fixed finite set $B \subset G$ we define its *impact function* by

$$\xi_B(m) = \xi_B(m, G) = \min\{|A + B| : A \subset G,\ |A| = m\}.$$

This is defined for all positive integers if G is infinite, and for $m \le |G|$ if G is finite.

This function embodies what can be told about cardinality of sumsets if one of the set is unrestricted up to cardinality. The name is a translation of Plünnecke's "Wirkungsfunktion", who first studied this concept systematically for density [9].

We will be interested mainly in the infinite case, and in this case the dependence on G can be omitted.

Lemma 3.1. *Let G, G' be infinite commutative groups, $G' \subset G$, and let $B \subset G'$ be a finite set. We have*

$$\xi_B(m, G) = \xi_B(m, G') \tag{3.1}$$

for all m.

Proof. Take an $A \subset G$, $|A| = m$ with $|A + B| = \xi_B(m, G)$. Let $A = A_1 \cup \cdots \cup A_k$ be its decomposition according to cosets of G'. For each $1 \le i \le k$ take an element x_i from the coset containing A_i so that the sets $A_i - x_i$ are pairwise disjoint; this is easily done as long as G' is infinite. The set

$$A' = \bigcup (A_i - x_i)$$

satisfies $A' \subset G'$, $|A'| = n$ and

$$|A' + B| \le \sum |A_i - x_i + B| = \sum |A_i + B| = |A + B| = \xi_B(m, G),$$

hence $\xi_B(m, G') \le \xi_B(m, G)$. The inequality in the other direction is obvious. □

In the case of finite groups the connection between $\xi_B(m, G)$ and $\xi_B(m, G')$ can also be described by arguments like in chapters 3 and 4 of Plünnecke's above mentioned book [9]. We restrict our attention to infinite groups, and henceforth omit the reference to G and write just $\xi_B(m)$ instead.

Let G be a torsionfree group. Take a finite $B \subset G$, and let G' be the subgroup generated by $B - B$, that is, the smallest subgroup such that B is contained in a single coset. Let $B' = B - a$ with some $a \in B$, so that $B' \subset G'$. The group G', as any finitely generated torsionfree group, is isomorphic to the additive group $\mathbb{Z}^d$ for some d. Let $\varphi : G' \to \mathbb{Z}^d$ be such an isomorphism and $B'' = \varphi(B')$. By Lemma 3.1 we have

$$\xi_B = \xi_{B'} = \xi_{B''},$$

so when studying the impact function we can restrict our attention to sets in $\mathbb{Z}^d$ that contain the origin and generate the whole lattice; we then study the set "in its natural habitat".

Definition 3.2. Let B be a finite set in a torsionfree group G. By the *dimension* of B we mean the number d defined above, and denote it by $\dim B$. By the *hull volume* of B we mean the volume of the convex hull of the set B'' described above and denote it by $\operatorname{hv} B$.

The set B'' is determined up to an automorphism of $\mathbb{Z}^d$. These automorphisms are exactly linear maps of determinant ± 1, hence the hull volume is uniquely defined.

Observe that this dimension is not the same as the dimension described in the Introduction; in the case when $B \subset \mathbb{R}^k$ with some k, this is its dimension over the field of rationals.

Theorem 3.3. *Let B be a finite set in a torsionfree group G, $d = \dim B$, $v = \operatorname{hv} B$. We have*

$$\lim |kB|k^{-d} = v.$$

A proof can be found in [12], Section 11, though this form is not explicitly stated there. An outline is as follows. By using the arguments above we may assume that $B \subset \mathbb{Z}^d$, $0 \in B$ and B generates $\mathbb{Z}^d$. Let B^* be the convex hull of B. Then kB is contained in $k \cdot B^*$. The number of lattice points in $k \cdot B$ is asymptotically $\mu(k \cdot B^*) = k^d v$; this yields an upper estimate. To get a lower estimate one proves that with some constant p, kB contains all the lattice points inside translate of $(k-p) \cdot B^*$; this is Lemma 11.2 of [12].

This means that the hull volume can be defined without any reference to convexity and measure, and this definition can even be extended to commutative semigroups. This follows from the following result of Khovanskii [5], [6]; for a simple proof see [8].

Theorem 3.4 (Khovanskii). *Let B be a finite set in a commutative semigroup. There is a k_0, depending on the set B, such that $|kB|$ is a polynomial function of k for $k > k_0$.*

Definition 3.5. Let B be a finite set in a commutative semigroup, and let vk^d be the leading term of the polynomial which coincides with $|kB|$ for large k. By the *dimension* of B we mean the degree d of this polynomial, and by the *hull volume* we mean the leading coefficient v.

It turns out that in $\mathbb{Z}^d$, hence in any torsionfree group, the dimension and hull volume determine the asymptotic behaviour of the impact function.

Theorem 3.6. *Let B be a finite set in a torsionfree commutative group G, $d = \dim B$, $v = \operatorname{hv} B$. We have*

$$\lim(\xi_B(m)^{1/d} - m^{1/d}) = v^{1/d}.$$

This is the main result (Theorem 3.1) of [12]. In the same paper I announce the same result for non necessarily torsionfree commutative groups without proof (Theorem 3.4). In a general semigroup $A+B$ may consist of a single element, so an attempt to an immediate generalization fails.

Problem 3.7. Does the limit $\lim \xi_B(m)^{1/d} - m^{1/d}$ exist in general commutative semigroups? Is there a condition weaker than cancellativity to guarantee its positivity?

Theorem 3.6 can be effectivized as follows (Theorems 3.2 and 3.3 of [12]).

Theorem 3.8. *With the notations of the previous theorem, if $d \geq 2$ and $m \geq v$, we have*

$$\xi_B(m) \leq m + dv^{1/d}m^{1-1/d} + c_1 v^{2/d}m^{1-2/d},$$
$$\xi_B(m)^{1/d} - m^{1/d} \leq v^{1/d} + c_2 v^{2/d}m^{-1/d}.$$

(c_1, c_2 depend on d.) With $n = |B|$ for large m we have

$$\xi_B(m) \geq m + dv^{1/d}m^{1-1/d} - c_3 v^{\frac{d+3}{2d}} n^{-1/2} m^{1-\frac{3}{2d}},$$

$$\xi_B(m)^{1/d} - m^{1/d} \geq v^{1/d} - c_4 v^{\frac{d+3}{2d}} n^{-1/2} m^{-1/(2d)}.$$

Probably the real error terms are much smaller than these estimates. For $d = 1$ we have the obvious inequality $\xi_B(m) \leq m + v$, with equality for large m because the integers $\xi_B(m) - m$ cannot converge to v otherwise. For $d = 2$ already $\sqrt{\xi_B(m)} - \sqrt{m}$ can converge to $\sqrt{v}$ from both directions.

Theorem 3.9. *The impact function of the set $B = \{0, e_1, e_2\} \subset \mathbb{Z}^2$ satisfies*

$$\sqrt{\xi_B(m)} - \sqrt{m} > \sqrt{v} \tag{3.2}$$

for all m.

The impact function of the set $B = \{0, e_1, e_2, -(e_1+e_2)\} \subset \mathbb{Z}^2$ satisfies

$$\sqrt{\xi_B(m)} - \sqrt{m} < \sqrt{v} \tag{3.3}$$

for infinitely many m.

Inequality (3.2) was announced in [12] without proof as Theorem 4.1, and it is a special case of Gardner and Gronchi's Theorem 2.6. Inequality 3.3 is Theorem 4.3 of [12].

I cannot decide whether there is a set such that $\sqrt{\xi_B(m)} - \sqrt{m} < \sqrt{v}$ for all m.

4. The impact volume

Besides cardinality we saw the hull volume as a contender for the title "discrete volume". For both we had something resembling the Brunn–Minkowski inequality;

for cardinality we had Gardner and Gronchi's Theorem 2.6, which has the (necessary) factor $d!$, and for the hull volume we have Theorem 3.6, which only holds asymptotically.

There is an easy way to find a quantity for which the analogue of the Brunn–Minkowski inequality hods exactly: we can make it a definition.

Definition 4.1. The d-dimensional *impact volume* of a set B (in an arbitrarily commutative group) is the quantity

$$\mathrm{iv}_d(B) = \inf_{m \in \mathbb{N}} \left(\xi_B(m)^{1/d} - m^{1/d}\right)^d.$$

Note that the d above may differ from the dimension of B, in fact, it need not be an integer. It seems, however, that the only really interesting case is $d = \dim B$.

The following statement list some immediate consequences of this definition.

Statement 4.2. *Let B be a finite set in a commutative torsionfree group.*

(a) $\mathrm{iv}_d(B)$ *is a decreasing function of d.*

(b) *If $|B| = n$, then*

$$\mathrm{iv}_1(B) = n - 1$$

and

$$\mathrm{iv}_d(B) \leq \left(n^{1/d} - 1\right)^d \tag{4.1}$$

for every d.

(c) $\mathrm{iv}_d(B) = 0$ *for* $d > \dim B$.

(d) *For every pair A, B of finite sets in the same group and every d we have*

$$\mathrm{iv}_d(A + B)^{1/d} \geq \mathrm{iv}_d(A)^{1/d} + \mathrm{iv}_d(B)^{1/d}. \tag{4.2}$$

The price we have to pay for the discrete Brunn–Minkowski inequality (4.2) is that there is no easy way to compute the impact volume for a general set. We have the following estimates.

Theorem 4.3. *Let B be a finite set in a commutative torsionfree group,* $\dim B = d$*, $|B| = n$. We have*

$$\binom{n-d}{d!} \leq \mathrm{iv}_d(B) \leq \mathrm{hv}\, B, \tag{4.3}$$

with equality in both places if B is a long simplex.

The first inequality follows form Theorem 2.6 of Gardner and Gronchi, the second from Theorem 3.6.

Problem 4.4. What is the *maximal* possible value of $\mathrm{iv}_d(B)$ for n-element d-dimensional sets? Is perhaps the bound in (4.1) exact?

We now describe the impact volume for another important class of sets, namely cubes.

Theorem 4.5. *Let $n_1, \dots, n_d$ be positive integers and let*

$$B = \{(x_1, \dots, x_d) \in \mathbb{Z}^d : 0 \le x_i \le n_i\}. \tag{4.4}$$

We have

$$\mathrm{iv}_d(B) = \mathrm{hv}\, B = v = n_1 \dots n_d.$$

Problem 4.6. Is is true that when B is the set of lattice points within a convex lattice polytope, then $\mathrm{hv}\, B$ and $\mathrm{iv}_d(B)$ are very near?

They may differ, as the second example in Theorem 3.9 shows.

We shall deduce Theorem 4.5 from the following one.

Theorem 4.7. *Let $G = G_1 \times G_2$ be a commutative group represented as the direct product of the groups G_1 and G_2. Let $B = B_1 \times B_2 \subset G$ be a finite set with $B_1 \subset G_1$, $B_2 \subset G_2$. We have*

$$\mathrm{iv}_d(B) \ge \mathrm{iv}_{d-1}(B_1)\mathrm{iv}_1(B_2). \tag{4.5}$$

Proof. Write $\mathrm{iv}_d(B) = v$, $\mathrm{iv}_{d-1}(B_1) = v_1$, $\mathrm{iv}_1(B_2) = v_2$ (which is $= |B| - 1$ if G_2 is torsionfree). We want to estimate $|A + B|$ from below for a general set $A \subset G$ with $|A| = m$.

Firs we transform them to some standard form; this will be the procedure what Gardner and Gronchi call compression. Let A_1 be the projection of A to G_1, and for an $x \in A_1$ write

$$A(x) = \{y \in G_2 : (x, y) \in A\}.$$

Let

$$A' = \{(x, i) : x \in A_1, i \in \mathbb{Z},\ 0 \le i \le |A(x)| - 1\}$$

and

$$B' = \{(x, i) : x \in B_1, i \in \mathbb{Z},\ 0 \le i \le v_2\}.$$

We have $A', B' \subset G' = G_1 \times \mathbb{Z}$.

Lemma 4.8. *We have*

$$|A'| = |A|, \quad |A' + B'| \le |A + B|. \tag{4.6}$$

Proof. The equality is clear. To prove the inequality, write $S = A + B$, $S' = A' + B'$. With the obvious notation, we will show that

$$|S'(x)| \le |S(x)|$$

for each x. To this end observe that

$$S(x) = \bigcup_{x'+x''=x} (A(x') + B(x'')) = \bigcup_{x' \in x - B_1} A(x') + B_2,$$

hence

$$|S(x)| \geq \max_{x' \in x - B_1} |A(x') + B_2| \geq \max_{x' \in x - B_1} |A(x')| + v_2.$$

Similarly

$$S'(x) = \bigcup_{x'+x''=x} (A'(x') + B'(x'')) = \bigcup_{x' \in x - B_1} [0, |A(x')| + v_2 - 1],$$

and so

$$|S'(x)| = \max_{x' \in x - B_1} |A(x')| + v_2. \qquad \square$$

Now we continue the proof of the theorem. Decompose A' into layers according to the value of the second component; write

$$A' = \bigcup_{i=0}^{k} L_i \times \{i\},$$

where $k = \max |A(x)|$, $L_i \subset G_1$. Write $|L_i| = m_i$. We have $L_0 \supset L_1 \supset \cdots \supset L_k$, consequently $m_0 \geq m_1 \geq \cdots \geq m_k$.

The set S' is the union of the sets $(L_i + B_1) \times \{i + j\}$, $0 \leq i \leq v_2$. By the above inclusion it is sufficient to consider the L_i with the smallest possible i, that is,

$$S' = (L_0 + B_1) \times \{0, 1, \dots, v_2\} \cup \bigcup_{i=1}^{k} (L_i + B_1) \times \{i + v_2\}.$$

We obtain that

$$|S'| = v_2 |L_0 + B_1| + \sum_{i=0}^{k} |L_i + B_1|. \tag{4.7}$$

To estimate the summands we use the $d-1$-dimensional impact of B_1, and we get

$$|L_i + B_1| \geq \left(m_i^{\frac{1}{d-1}} + v_1^{\frac{1}{d-1}} \right)^{d-1} \geq \frac{m_i}{m_0} \left(m_0^{\frac{1}{d-1}} + v_1^{\frac{1}{d-1}} \right)^{d-1};$$

the second inequality follows from $m_i \leq m_0$. By substituting this into (4.7) and recalling that $\sum m_i = m$ we obtain

$$|S| \geq \left(v_2 + \frac{m}{m_0} \right) \left(m_0^{\frac{1}{d-1}} + v_1^{\frac{1}{d-1}} \right)^{d-1}. \tag{4.8}$$

Consider the right side as a function of the real variable m_0. By differentiating we find that it assumes its minimum at

$$m_0 = v_1^{1/d} (m/v_2)^{1-1/d}.$$

(This minimum typically is not attained; this m_0 may be < 1 or $> m$, and it is generally not integer). Substituting this value of m_0 into (4.8) we obtain the desired bound

$$|S| \geq \left(m^{1/d} + (v_1 v_2)^{1/d} \right)^d. \qquad \square$$

Problem 4.9. Does equality always hold in Theorem 4.7?

I expect a negative answer.

Problem 4.10. Can Theorem 4.7 be extended to an inequality of the form

$$\mathrm{iv}_{d_1+d_2}(B_1 \times B_2) \geq \mathrm{iv}_{d_1}(B_1)\mathrm{iv}_{d_2}(B_2)?$$

Proof of Theorem 4.5. To prove $\geq$ we use induction on d. The case $d = 1$ is obvious, and Theorem 4.7 provides the inductive step.

This means that with the cube B defined in (4.4) we have

$$|A + B| \geq \left(|A|^{1/d} + v^{1/d}\right)^d.$$

Equality can occur for infinitely many values of $|A|$, namely it holds whenever A is also a cube of the form

$$A = \{(x_1, \ldots, x_d) \in \mathbb{Z}^d : 0 \leq x_i \leq kn_i - 1\}$$

with some integer k; we have $|A| = k^d v$, $|A + B| = (k + 1)^d v$. It may be difficult to describe $\xi_B(m)$ for values of m which are not of the form $k^d v$. Possibly an argument like Gardner and Gronchi's for the simplex may work.

Observe that these special sets A are not homothetic to B; in particular, $A = B$ may not yield a case of equality. □

As Theorem 4.3 shows, the impact volume can be $d!$ times smaller than cardinality. The example we have of this phenomenon, the long simplex, is, however, "barely" d-dimensional, and I expect that a better estimates hold for a "substantially" d-dimensional set.

Definition 4.11. The *thickness* $\vartheta(B)$ of a set $B \subset \mathbb{R}^d$ is the smallest integer k with the property that there is a hyperplane P of $\mathbb{R}^d$ and $x_1, \ldots, x_k \in \mathbb{R}^d$ such that $B \subset \bigcup_{i=1}^k P + x_j$.

Conjecture 4.12. For every $\varepsilon > 0$ and d there is a k such that for every $B \subset \mathbb{R}^d$ with $\vartheta(B) > k$ we have $\mathrm{iv}_d(B) > (1 - \varepsilon)|B|$.

This conjecture would yield a discrete Brunn–Minkowski inequality of the form

$$|A + B|^{1/d} \geq |A|^{1/d} + (1 - \varepsilon)|B|^{1/d}$$

assuming a bound on the thickness of B. Such an inequality is true at least in the special case $A = B$. This can be deduced from a result of Freiman ([3], Lemma 2.12; see also Bilu [1]), which sounds as follows. If $A \subset \mathbb{R}^d$ and $|2A| < (2^d - \varepsilon)|A|$, then there is a hyperplane P such that $|P \cap A| > \delta|A|$, with $\delta = \delta(d, \varepsilon) > 0$.

5. Meditation on the continuous case

Let A, B be Borel sets in $\mathbb{R}^d$. The Brunn–Minkowski inequality (1.2) estimates $\mu(A + B)$ in a natural way, with equality if A and B are homothetic convex sets.

Like in the discrete case, we can define the *impact function* of the set B by

$$\xi_B(a) = \inf\{\mu(A + B) : \mu(A) = a\}.$$

Thus (1.2) is equivalent to

$$\xi_B(a) \geq \left(a^{1/d} + \mu(B)^{1/d}\right)^d,$$

and this is the best possible estimate in terms of $\mu(B)$ only.

To measure the degree of nonconvexity we propose to use the measure of the convex hull beside the measure of the set. This is analogous to the hull volume, and it is sufficient to describe the asymptotic behaviour of ξ.

Theorem 5.1 ([13], Theorem 1.). *For every bounded Borel set $B \subset \mathbb{R}^d$ of positive measure we have*

$$\lim_{a\to\infty} \xi_B(a)^{1/d} - a^{1/d} = \mu(\operatorname{conv} B)^{1/d}.$$

This is the continuous analogue of Theorem 3.6, and there is an analogue to the effective version Theorem 3.8 as well.

Note that by considering sets homothetic to conv B we immediately obtain

$$\xi_B(a)^{1/d} \leq a^{1/d} + \mu(\operatorname{conv} B)^{1/d},$$

thus we need only to give a lower estimate. This is as follows.

Theorem 5.2 ([13], Theorem 2.). *Let $\mu(B) = b$, $\mu(\operatorname{conv} B) = v$. We have*

$$\xi_B(a)^{1/d} \geq a^{1/d} + v^{1/d}\left(1 - c(v/b)^{1/2}(v/a)^{1/(2d)}\right)$$

$$\xi_B(a) \geq a + dv^{1/d}a^{1-1/d}\left(1 - c(v/b)^{1/2}(v/a)^{1/(2d)}\right)$$

with a suitable positive constant c depending on d.

If $v > b$, we get a nontrivial improvement over the Brunn–Minkowski inequality for $a > a_0(b, v)$. It would be desirable to find an improvement also for small values of a, or, even more, to find the best estimate in terms of $\mu(B)$ and $\mu(\operatorname{conv} B)$.

The exact bound and the structure of the extremal set may be complicated. This is already so in the case $d = 1$, which was solved in [10]. Observe that in one dimension $\mu(\operatorname{conv} B)$ is the diameter of B.

Theorem 5.3 ([10], Theorem 2). *Let $B \subset \mathbb{R}$, and write $\mu(B) = b$, $\mu(\operatorname{conv} B) = v$. If*

$$a \geq \frac{v(v-b)}{2b} + \frac{b\{v/b\}(1 - \{v/b\})}{2}, \tag{5.1}$$

then $\xi_B(a) = a + v$. *If* (5.1) *does not hold, then let k be the unique positive integer satisfying*

$$\frac{k(k-1)}{2} \le \frac{a}{b} < \frac{k(k+1)}{2}$$

and define δ *by*

$$\frac{a}{b} = \frac{k(k-1)}{2} + \delta k.$$

We have

$$\xi_B(a) \ge a + (k+\delta)b,$$

and equality holds if $B = [0, b] \cup \{v\}$.

A set A such that $\xi_B(a) = \mu(A+B)$ for the above set B is given by

$$A = [0, (k-1+\delta)b] \cup [v, v + (k-2+\delta)b] \cup \cdots \cup [(k-1)v, (k-1)v + \delta b].$$

A less exact, but simple and still quite good lower bound sounds as follows.

Corollary 5.4 ([10], Theorem 1). *Let* $B \subset \mathbb{R}$, *and write* $\mu(B) = b$, $\mu(\operatorname{conv} B) = v$. *We have*

$$\xi_B(a) \ge \min\bigl(a + v, (\sqrt{a} + \sqrt{b/2})^2\bigr).$$

A comparison with the 2-dimensional Brunn–Minkowski inequality gives the following interpretation: initially a long one-dimensional set B tries to behave as if it were a two-dimensional set of area $b/2$.

It can be observed that (5.4) is weaker than the obvious inequality

$$\mu(A+B) \ge \mu(A) + \mu(B) \tag{5.2}$$

for small a. For small values of a Theorem 5.3 yields the following improvement of (5.2).

Corollary 5.5 ([10], Corollary 3.1). *If* $a \le b$, *then we have*

$$\mu(A+B) \ge \min(2a + b, a + v).$$

If $b < a \le 3b$, *then we have*

$$\mu(A+B) \ge \min\left(\frac{3}{2}(a+b), a + v\right).$$

Problem 5.6. How large must $\mu(A+B)$ be if $\mu(A)$, $\mu(B)$, $\mu(\operatorname{conv} A)$ and $\mu(\operatorname{conv} B)$ are given?

What are the minima of $\mu(A+A)$ and $\mu(A-A)$ for fixed $\mu(A)$ and $\mu(\operatorname{conv} A)$?

The results above show that for $d = 1$ (like in the discrete case, but for less obvious reasons) the limit relation becomes an equality for $a > a_0$. Again, this is no longer the case for $d = 2$.

An example of a set $B \subset \mathbb{R}^2$ such that

$$\xi_B(a)^{1/2} < a^{1/2} + v^{1/2}$$

will hold for certain arbitrarily large values of a is as follows.

Let $0 < c < 1$ and let B consist of the square $[0, c] \times [0, c]$ and the points $(0, 1)$, $(1, 0)$ and $(1, 1)$. Hence $b = c^2$ and $v = 1$.

For an integer $n \geq 1$ put

$$A_n = [0, n] \times [0, n] \cup \bigcup_{j=0}^{n} [j, j + c] \times [n, n + c] \cup \bigcup_{j=0}^{n-1} [n, n + c] \times [j, j + c].$$

Thus A_n consists of a square of side n and $2n + 1$ small squares of side c, hence

$$\mu(A_n) = n^2 + (2n + 1)b.$$

We can easily see that $A_n + B = A_{n+1}$. Hence by considering the set $A = A_n$ we see that for a number a of the form $a = n^2 + (2n + 1)b$ we have

$$\xi_B(a) \leq \mu(A_{n+1}) = (n + 1)^2 + (2n + 3)b < \left(\sqrt{a} + 1\right)^2.$$

A more detailed calculation leads to

$$\xi_B(a)^{1/2} \leq a^{1/2} + 1 - ca^{-1}$$

(for these special values of a).

If we tried to define an impact volume in the continuous case, we would recover the volume, at least for compact sets. Still, the above results and questions suggest that ordinary volume is not the best tool to understand additive properties. Perhaps one could try to modify the definition of impact volume by requiring $\mu(A) \geq \mu(B)$. So put

$$\mathrm{iv}_*(B) = \inf_{a \geq \mu(B)} \left(\xi_B(a)^{1/d} - a^{1/d}\right)^d.$$

Problem 5.7. Find a lower estimate for $\mathrm{iv}_*(B)$ in terms of $\mu(B)$ and $\mu(\mathrm{conv}\, B)$.

6. Back to one dimension

The results in the previous section, Theorem 5.3 and Corollaries 5.4 and 5.5 show that one can have nontrivial results in the seemingly uninteresting one-dimensional case. We now try to do the same, and will find bounds on $|A + B|$ using the cardinality and hull volume of B. Observe that in one dimension the hull volume is the smallest l such that B is contained in an arithmetic progression $\{b, b+q, \dots, b+lq\}$: the *reduced diameter* of B.

It is possible to give bounds using nothing else than the hull volume.

Theorem 6.1. *Let B be a one-dimensional set in a torsionfree commutative group,* $\operatorname{hv} B = v \geq 3$.

(a) *For*

$$m > \frac{(v-1)(v-2)}{2}$$

we have $\xi_B(m) = m + v$.

(b) *If*

$$\frac{(k-1)(k-2)}{2} < m \leq \frac{k(k-1)}{2}$$

with some integer $2 \leq k < v$, then $\xi_B(m) \geq m + k$. Equality holds for the set $B = \{0, 1, v\} \subset \mathbb{Z}$.

For $v \leq 2$ we have obviously $\xi_B(m) = m + v$ for all m (such a set cannot be anything else than a $v+1$-term arithmetic progression).

This will be deduced from the following result, where the cardinality of B is also taken into account.

Theorem 6.2. *Let B be a one-dimensional set in a torsionfree commutative group,* $\operatorname{hv} B = v \geq 3$, $|B| = n$. *Define w by*

$$w = \min_{d|v,\, d \leq n-2} d \left\lceil \frac{n-2}{d} \right\rceil. \tag{6.1}$$

For every m we have

$$\xi_B(m) \geq m + \min\left(v, \frac{w}{2} + \min_{t \in \mathbb{N}} \left(\frac{m}{t} + \frac{tw}{2}\right)\right). \tag{6.2}$$

The minimum is attained either at the floor or at the ceiling of $\sqrt{2m/w}$. Unlike the previous theorem, typically we do not have examples of equality, and the extremal value and the structure of extremal sets is probably complicated. Also the value of w depends on divisibility properties of v and n. After the proof we give a less exact but simpler corollary.

Proof. By Lemma 3.1 we may assume that $B \subset \mathbb{Z}$, its smallest element is 0 and it generates $\mathbb{Z}$; then its largest element is just v.

Lemma 6.3. *Let B' be the set of residues of elements of B modulo v. For every nonempty $X \subset \mathbb{Z}_v$ we have*

$$|X + B'| \geq \min(|X| + w, v). \tag{6.3}$$

Proof. By Kneser's theorem we have

$$|X + B'| \geq |X + H| + |B' + H| - |H|$$

with some subgroup H of the additive group $\mathbb{Z}_v$. Write $|H| = d$; clearly $d|v$. If $d = v$, we have $|X + H| = v$ and we are ready. Assume $d < v$. B' contains 0 and it generates $\mathbb{Z}_v$, hence it cannot be contained in H so we have $|B' + H| \geq 2|H| = 2d$. This gives the desired bound if $d > n - 2$. Assume $d \leq n - 2$. Since $|B' + H|$ is a multiple of d and it is at least $|B'| = n - 1$, we obtain

$$|B' + H| \geq d \left\lceil \frac{n-1}{d} \right\rceil = d\left(1 + \left[\frac{n-2}{d}\right]\right) \geq d + w. \qquad \square$$

We resume the proof of Theorem 6.2. Take a set $A \subset \mathbb{Z}$, $|A| = m$. We are going to estimate $|A + B|$ from below.

For $j \in \mathbb{Z}_v$ let $u(j)$ be the number of integers $a \in A$, $a \equiv j \pmod v$ and let $U(j)$ be the corresponding number for the sumset $A + B$. We have

$$U(j) \geq u(j) + 1 \tag{6.4}$$

whenever $U(j) > 0$; this follows by adding the numbers $0, v$ to each element of A in this residue class if $u(j) > 0$, and holds obviously for $u(j) = 0$. We also have

$$U(j) \geq u(j - b) \tag{6.5}$$

for every $b \in B'$. Write

$$r(k) = \{j : u(j) \geq k\},$$

$$R(k) = \{j : U(j) \geq k\}.$$

Inequality (6.4) implies

$$R(k) \supset r(k - 1) \quad (k \geq 2), \tag{6.6}$$

and inequality (6.5) implies

$$R(k) \supset r(k) + B' \quad (k \geq 1). \tag{6.7}$$

First case. $U(j) > 0$ for all j. In this case by summing (6.4) we get

$$|A + B| = \sum U(j) \geq v + \sum u(j) = |A| + v.$$

Second case. There is a j with $U(j) = 0$. Then we have $|R(k)| < v$ for every $k > 0$. An application of Lemma 6.3 to the sets $r(k)$ yields, in view of (6.7),

$$|R(k)| \geq |r(k)| + w \tag{6.8}$$

as long as $r(k) \neq \emptyset$. Let t be the largest integer with $r(t) \neq \emptyset$. We have (6.8) for $1 \leq k \leq t$, and (6.6) yields

$$|R(k)| \geq |r(k-1)| \tag{6.9}$$

for all $k \geq 2$. Consequently for $1 \leq k \leq t+1$ we have

$$|R(k)| \geq \frac{k-1}{t}|r(k-1)| + \left(1 - \frac{k-1}{t}\right)(|r(k)| + w). \tag{6.10}$$

Indeed, for $k = 1$ (6.10) is identical with (6.8), for $k = t+1$ it is identical with (6.9) and for $2 \leq k \leq t$ it is a linear combination of the two.

By summing (6.10) we obtain

$$\begin{aligned} |A+B| &= \sum_{k \geq 1} |R(k)| \geq \sum_{k=1}^{t+1} |R(k)| \geq \frac{t+1}{2} w + \left(1 + \frac{1}{t}\right) \sum_{k=1}^{t} |r(k)| \\ &= \frac{t+1}{2} w + \left(1 + \frac{1}{t}\right)|A|, \end{aligned}$$

as claimed in (6.2). □

Corollary 6.4. *With the assumptions and notations of Theorem* 6.2 *we have*

$$\xi_B(m) \geq \min\left(m + v, \left(\sqrt{m} + \sqrt{w/2}\right)^2\right). \tag{6.11}$$

Proof. This follows from (6.2) and the inequality of arithmetic and geometric means. □

This can be interpreted as that the set tries to imitate a two-dimensional set of area $w/2$.

Proof of Theorem 6.1. Parts (a)–(b) of the theorem can be reformulated as follows: if $\xi_B(n) \leq m + k$ with some $k < v$, then $m \leq k(k-1)/2$. Theorem 6.2 yields (using only that $w \geq 1$) the existence of a positive integer t such that

$$\frac{m}{t} + \frac{t+1}{2} \leq k,$$

hence

$$m \leq kt + \frac{t(t+1)}{2}.$$

The right side, as a function of t, is increasing up to $k - 1/2$ and decreasing afterwards; the minimal values at integers are assumed at $t = k-1$ and k, and both are equal to $k(k-1)/2$.

To show the case of equality in case (b), write $m = k(k-1)/2 - l$ with $0 \leq l \leq k-2$. The set A will contain the integers in the intervals $[iv, iv + k - 3 - i]$ for $0 \leq i \leq l - 1$ and $[iv, iv + k - 2 - i]$ for $l \leq i \leq k - 2$. □

We illustrate the strength of Theorem 6.2 by deducing from it the two-dimensional estimate

$$\xi_L(m) > \big(\sqrt{m} + \sqrt{(n-2)/2}\big)^2$$

for the long triangle $L = L_{2n}$. Indeed, a suitable linear mapping maps this set L onto the set $B = \{0, 1, \dots, n-2, v\}$ with arbitrary v. If we choose v to be prime, then in (6.1) we have $w = n-2$, and if v is so large that $m + l > \big(\sqrt{m} + \sqrt{w}\big)^2$, then from Corollary 6.4 we obtain

$$\xi_L(m) \geq \xi_B(m) \geq \big(\sqrt{m} + \sqrt{w/2}\big)^2.$$

This is essentially the two-dimensional case of Theorem 2.6 of Gardner and Gronchi.

On the other hand, for small values of m this inequality is weak, can even be worse than the obvious bound $|A+B| \geq |A| + |B| - 1$. There are results that are especially suited to the study of small values; we quote two of them. In both let $A, B \subset \mathbb{Z}$, $A = \{a_1, \dots, a_m\}$, $B = \{b_1, \dots, b_n\}$ with $0 = a_1 < \cdots < a_m = u$, $0 = b_1 < \cdots < b_n = v$.

Theorem 6.5 (Freiman [2]). *If* $\gcd(a_1, \dots, a_m, b_1, \dots, b_n) = 1$ *and* $u \leq v$, *then*

$$|A+B| \geq \min(m+v, m+n+\min(m,n)-3).$$

This bears a remarkable similarity to the two-dimensional case of Theorem 1.2 (and it can be deduced like Theorem 2.6)

Theorem 6.6 (Lev and Smelianski [7]). *If* $\gcd(b_1, \dots, b_n) = 1$ *and* $u \leq v$, *then*

$$|A+B| \geq \min(m+v, n+2m-\delta),$$

where $\delta = 3$ *if* $u = v$ *and* $\delta = 2$ *if* $u < v$.

Observe that the above theorems cannot be directly compared to ours because of the somewhat different structure of the assumptions.

Problem 6.7. Find a common generalization of Theorems 6.2 and 6.6.

References

[1] Bilu, Y., Structure of sets with small sumset. Structure theory of set addition. *Astérisque* **258** (1999), 77–108.

[2] Freiman, G., Inverse problems of additive number theory. VI. On the addition of finite sets. III. *Izv. Vyss. Ucebn. Zaved. Matematika* **3** (28) (1962), 151–157 (in Russian).

[3] —, *Foundations of a structural theory of set addition*. Transl. Math. Monogr. 37, Amer. Math. Soc., Providence, RI, 1973.

[4] Gardner, R. J., and Gronchi, P., A Brunn-Minkowski inequality for the integer lattice. *Trans. Amer. Math. Soc.* **353** (2001), 3995–4024.

[5] Khovanskii, A. G., Newton polyhedron, Hilbert polynomial, and sums of finite sets. *Funct. Anal. Appl.* **26** (1992), 276–281.

[6] —, Sums of finite sets, orbits of commutative semigroups, and hilbert functions. *Funct. Anal. Appl.* **29** (1995), 102–112.

[7] Lev, V. F., and Smeliansky, P., On addition of two distinct sets of integers. *Acta Arith.* **70** (1995), 85–91.

[8] Nathanson, M. B., and Ruzsa, I. Z., Polynomial growth of sumsets in abelian semigroups. *J. Théor. Nombres Bordeaux* **14** (2002), 553–560.

[9] Plünnecke, H., *Eigenschaften und Abschätzungen von Wirkungsfunktionen*. Gesellschaft für Mathematik und Darenverarbeitung, Bonn, 1969.

[10] Ruzsa, I. Z., Diameter of sets and measure of sumsets. *Monatsh. Math.* **112** (1991), 323–328.

[11] —, Sum of sets in several dimensions. *Combinatorica* **14** (1994), 485–490.

[12] —, Sets of sums and commutative graphs. *Studia Sci. Math. Hungar.* **30** (1–2) (1995), 127–148.

[13] —, The Brunn-Minkowski inequality and nonconvex sets. *Geom. Dedicata* **67** (1997), 337–348.

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