Combinatorics, Probability and Computing (2008) 17, 681–688.
doi:10.1017/S0963548308009139 Printed in the United Kingdom

Projecting Difference Sets on the Positive Orthant

RON HOLZMAN[1], VSEVOLOD F. LEV[2] and ROM PINCHASI[1†]

[1]Department of Mathematics, Technion-Israel Institute of Technology, Haifa 32000, Israel
(e-mail: holzman@techunix.technion.ac.il, room@math.technion.ac.il)

[2]Department of Mathematics, The University of Haifa at Oranim, Tivon 36006, Israel
(e-mail: seva@math.haifa.ac.il)

Received 25 January 2007; revised 1 April 2008; first published online 13 June 2008

Let $n \geqslant 1$ be an integer. Given a vector $a = (a_1, \dots, a_n) \in \mathbb{R}^n$, write

$$a^+ := (\max\{a_1, 0\}, \dots, \max\{a_n, 0\})$$

(the 'projection of a onto the positive orthant'). For a set $A \subseteq \mathbb{R}^n$ put $A^+ := \{a^+ : a \in A\}$ and $A - A := \{a - b : a, b \in A\}$. Improving previously known bounds, we show that $|(A-A)^+| \geqslant |A|^{3/5}/6$ for any finite set $A \subseteq \mathbb{R}^3$, and that $|(A-A)^+| \geqslant c|A|^{6/11}/(\log|A|)^{2/11}$ with an absolute constant $c > 0$ for any finite set $A \subseteq \mathbb{R}^4$ such that $|A| \geqslant 2$.

1. Introduction

In connection with a well-known conjecture by Graham (solved by Balasubramanian and Soundararajan in [1]), Granville and Roesler raised in [2] the following problem: for a finite set A of positive integers with prescribed cardinality $|A|$, what is the smallest possible size of the set

$$\left\{\tfrac{a}{\gcd(a,b)} : a, b \in A\right\} ?$$

As indicated in [2], this question can be equivalently re-stated as a combinatorial geometry problem, closely resembling the famous Erdős distance problem. Specifically, for a real number x let $x^+ := \max\{x, 0\}$. Given an integer $n \geqslant 1$, for a vector $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ write

$$a^+ := (a_1^+, \dots, a_n^+),$$

and for a set $A \subseteq \mathbb{R}^n$ put

$$A^+ := \{a^+ : a \in A\}$$

and

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

How small can $|(A-A)^+|$ be for a finite set $A \subseteq \mathbb{R}^n$ of the given size $|A|$?

† This author's research was supported by the Israel Science Foundation (grant No. 938/06).

There are two notable special cases where the inequality

$$|(A-A)^+| \geqslant |A| \tag{1.1}$$

holds true. First, it is easily seen to hold for $n = 1$; that is, when A is a set of real numbers. Indeed, (1.1) is sharp in this case: equality is attained if A is an arithmetic progression. Second, (1.1) holds if $A \subseteq \{0,1\}^n$; in this case the elements of A can be associated with subsets of an n-element set, and $(A-A)^+$ corresponds then to the family of all set-theoretic differences of these subsets. In this form the problem was studied in [3] by Marica and Schönheim. Their result says that the number of differences is at least as large as the number of subsets, which is equivalent to (1.1). Here, too, the estimate is sharp, as it follows by considering a family of sets that is closed under taking subsets.

Though (1.1) fails in general, it was observed by Granville and Roesler [2] and independently by Alon (personal communication) that the simple universal bound $|(A-A)^+| \geqslant \sqrt{|A|}$ is true for any dimension n and any finite set $A \subseteq \mathbb{R}^n$. This follows from the identity

$$x = x^+ - (-x)^+, \quad x \in \mathbb{R},$$

which shows that $a-b = (a-b)^+ - (b-a)^+$ for any $a, b \in A$, leading to $|A-A| \leqslant |(A-A)^+|^2$ and immediately implying the estimate in question.

For $n = 2$ a refined bound was obtained in [2].

Theorem 1.1 ([2]). *If $A \subseteq \mathbb{R}^2$ is finite, then $|(A-A)^+| \geqslant (|A|/2)^{2/3}$.*

The following example due to Freiman and the second-named author (but see [2]) shows that the estimate of Theorem 1.1 is of the best-possible order: if m is a positive integer and A is the set of integer vectors (x_1, x_2) with positive coordinates, satisfying $|(x_1+x_2) - (2m)^{2/3}/2| < (2m)^{1/3}/2$, then $|A| = m(1+o(1))$ and $|(A-A)^+| = \frac{3}{2}(2m)^{2/3}(1+o(1))$.

Clearly, the set just described can be embedded in $\mathbb{R}^n$ for any dimension $n \geqslant 2$, which shows that

$$\sqrt{m} \leqslant \min_{A \subseteq \mathbb{R}^n \,:\, |A|=m} |(A-A)^+| \leqslant \frac{3}{2}(2m)^{2/3}(1+o(1))$$

as $m \to \infty$, uniformly in n. A challenging open problem is to close the gap between the bounds, or at least to determine whether there exist positive absolute constants c and δ such that $|(A-A)^+| \geqslant c|A|^{0.5+\delta}$ holds for any $n \geqslant 1$ and any finite set $A \subseteq \mathbb{R}^n$. By all we know, it is even possible that $\delta = 1/6$ will do.

In this paper we improve the 'universal bound' for 3-dimensional and 4-dimensional sets.

Theorem 1.2. *If $A \subseteq \mathbb{R}^3$ is finite, then $|(A-A)^+| \geqslant \frac{1}{6}|A|^{3/5}$.*

Theorem 1.3. *There is an absolute constant $c > 0$ such that if $A \subseteq \mathbb{R}^4$ is finite and $|A| \geqslant 2$, then $|(A-A)^+| \geqslant c|A|^{6/11}/(\log|A|)^{2/11}$.*

We notice that our proof of Theorem 1.2 actually leads to a factor slightly larger than $1/6$, and that further minor improvements are possible without modifying the proof substantially. The

proof of Theorem 1.3 allows one to explicitly compute the constant c appearing in the statement of the theorem.

Theorems 1.2 and 1.3 are proved in Sections 3 and 4, respectively. An important auxiliary estimate, relating the size of $(A-A)^+$ with those of the projections of A on the coordinate hyperplanes, is established in the next section.

2. Positive differences and projections

Let $n \geqslant 1$ be an integer. For $i \in [1,n]$ we let

$$P_i := \{(x_1,\dots,x_n) \in \mathbb{R}^n : x_i = 0\}$$

(the ith coordinate hyperplane), and

$$L_i := \{(x_1,\dots,x_n) \in \mathbb{R}^n : x_j = 0 \quad \text{for } j \in [1,n],\ j \neq i\}$$

(the ith coordinate axis). Given a set $A \subseteq \mathbb{R}^n$, by A_i we denote the orthogonal projection of A on P_i.

We start with a very basic, but useful observation.

Lemma 2.1. *If $n \geqslant 2$ is an integer and $A \subseteq \mathbb{R}^n$ is finite, then $|(A-A)^+| \geqslant |(A_i - A_i)^+|$ holds for any $i \in [1,n]$.*

Proof. Since $(B_i)^+ = (B^+)_i$ for any $B \subseteq \mathbb{R}^n$, we have

$$|(A_i - A_i)^+| = |((A-A)_i)^+| = |((A-A)^+)_i| \leqslant |(A-A)^+|. \qquad \square$$

The next lemma is our main tool.

Lemma 2.2. *If $n \geqslant 2$ is an integer and $A \subseteq \mathbb{R}^n$ is a non-empty finite set, then*

$$|(A-A)^+| \geqslant \frac{1}{n^n} \frac{|A|^n}{|A_1| \cdots |A_n|}.$$

Proof. For each $i \in [1,n]$ we set $k_i := \lceil |A|/(n|A_i|) \rceil - 1$. There are $|A_n|$ lines in $\mathbb{R}^n$, parallel to L_n and passing through a point of A. On each of these lines we find k_n points of A with the largest nth coordinate and remove them from A, passing to a new set $A^{(n-1)} \subseteq \mathbb{R}^n$. (If a line contains fewer than k_n points of A, we remove them all.) The number of removed points is at most $k_n|A_n|$, and hence $|A^{(n-1)}| > (1-1/n)|A|$. Now we consider the lines in $\mathbb{R}^n$, parallel to L_{n-1} and passing through the points of $A^{(n-1)}$, and remove from $A^{(n-1)}$ those points which are among the k_{n-1} 'highest' on their lines. This yields a new set $A^{(n-2)}$, and since the total number of lines, considered at the second step, is at most $|A_{n-1}|$, we have $|A^{(n-2)}| > (1-2/n)|A|$. Repeating this procedure, after n steps we get a non-empty set $A^{(0)} \subseteq A$. Fix arbitrarily $a_0 \in A^{(0)}$. Since a_0 was not removed at the nth step, there are at least k_1+1 points in $A^{(1)}$ which coincide with a_0 in all but the first coordinate, and whose first coordinate is at least as large as that of a_0. Similarly, to each of these k_1+1 points in $A^{(1)}$ there correspond at least k_2+1 points in $A^{(2)}$ which coincide with the original point in all coordinates but the second one, and whose second coordinates are

greater than or equal to that of the original point. Continuing in this way, we eventually find at least

$$(k_1+1)\cdots(k_n+1) \geqslant \frac{1}{n^n}\,\frac{|A|^n}{|A_1|\cdots|A_n|}$$

points in A which are greater than or equal to a_0 in each coordinate. Subtracting a_0 from each of these points we obtain the required number of pairwise distinct points in $A-A$ with non-negative coordinates, and the result follows. □

We remark that the proof of Lemma 2.2 can easily be modified to show that any finite set $A \subseteq \mathbb{R}^n$ contains $\Omega\big(\frac{|A|^{n+1}}{|A_1|\cdots|A_n|}\big)$ comparable pairs of points (that is, pairs (a,b) such that all coordinates of $a-b$ are of the same sign).

Notice also that using Lemma 2.2 one can easily obtain an alternative proof of Theorem 1.1. Namely, if $A \subseteq \mathbb{R}^2$ is finite and non-empty, then $|(A-A)^+| \geqslant |A|^2/(4|A_1||A_2|)$ by Lemma 2.2, and furthermore $|(A-A)^+| \geqslant |A_1|$ and $|(A-A)^+| \geqslant |A_2|$ by Lemma 2.1. Multiplying out these estimates and extracting the cube root, we get $|(A-A)^+| \geqslant (|A|/2)^{2/3}$, as required.

3. 3-dimensional sets: proof of Theorem 1.2

To prove Theorem 1.2 we first translate A so that $0 \in A$ and all points of A have non-negative third coordinate. There is then a coordinate octant $\mathcal{O}$, containing at least $|A|/4$ points of A. Choose a two-element set $\{i,j\} \subseteq [1,3]$ so that for every point $(x_1,x_2,x_3) \in \mathcal{O}$ we have $x_i x_j \geqslant 0$. Renumbering the coordinate axes and replacing A with $-A$, if necessary, we can ensure that $i=1$, $j=2$, and $x_1, x_2 \geqslant 0$ for any $(x_1,x_2,x_3) \in \mathcal{O}$. (Notice, that the assumption $0 \in A$ is not affected by these manipulations; on the other hand, the assumption that all points of A have non-negative third coordinate may not be valid any longer.)

Set $B := A \cap \mathcal{O}$, so that $|B| \geqslant |A|/4$. Since $0 \in B$ and every point of B has non-negative first two coordinates, using Lemma 2.1 we get

$$|(B-B)^+| \geqslant |(B_3-B_3)^+| \geqslant |B_3^+| = |B_3|. \tag{3.1}$$

Furthermore, by Lemma 2.1 and Theorem 1.1,

$$|(B-B)^+| \geqslant |(B_1-B_1)^+| \geqslant (|B_1|/2)^{2/3}, \tag{3.2}$$

and similarly

$$|(B-B)^+| \geqslant (|B_2|/2)^{2/3}. \tag{3.3}$$

Finally, by Lemma 2.2,

$$|(B-B)^+| \geqslant \frac{1}{27}\,\frac{|B|^3}{|B_1||B_2||B_3|}. \tag{3.4}$$

From (3.1)–(3.4) we obtain

$$|(B-B)^+|^5 \geqslant \frac{|B|^3}{4\cdot 27} \geqslant \frac{|A|^3}{4^4\cdot 27},$$

whence

$$|(A-A)^+| \geqslant |(B-B)^+| \geqslant \frac{|A|^{3/5}}{(6912)^{1/5}} \geqslant \frac{1}{6}\,|A|^{3/5}.$$

This proves Theorem 1.2.

4. 4-dimensional sets: proof of Theorem 1.3

The following general result will eventually be applied to the hypergraph whose vertices are lines through the points of A, parallel to the coordinate axes, and whose edges are quadruples of such lines meeting at the points of A. This will allow us to pass from A to a subset such that every line, parallel to a coordinate axis and intersecting this subset, actually contains 'many' points of the subset.

Lemma 4.1. *Let $k \geqslant 2$ be an integer and suppose that $G = (V, E)$ is a hypergraph, every edge of which is incident to at most k vertices. If G has at least two non-isolated vertices, then there exists an induced hypergraph $G' = (V', E')$ with $|E'| \geqslant 0.9|E|$ and such that the minimal degree of G' is at least $\gamma|E|/(|V'|\log|V'|)$, where $\gamma = 1/(10\log k)$.*

We present here an elegant proof, kindly communicated to us by Noga Alon.

Proof of Lemma 4.1. Dropping isolated vertices, we can assume that every vertex of G has degree at least 1. Let $n := |V|$ and $m := |E|$. If $m \leqslant \gamma^{-1} n\log n$, then we can take $G' = G$; assume that $m > \gamma^{-1} n\log n$.

If G has a vertex of degree smaller than $\gamma m/(n\log n)$, drop it, passing to an induced subgraph on $n-1$ vertices; if this subgraph has a vertex of degree smaller than $\gamma m/((n-1)\log(n-1))$, drop this vertex, passing to an induced subgraph on $n-2$ vertices, *etc.* We claim that the process ends before we get to a subgraph on $n_0 := \lfloor n^{1/k} \rfloor$ vertices. For, if this is wrong, then the number of edges we dropped to get to the subgraph on n_0 vertices is less than

$$\begin{aligned}
&\gamma m\left(\frac{1}{n\log n} + \frac{1}{(n-1)\log(n-1)} + \cdots + \frac{1}{(n_0+1)\log(n_0+1)}\right)\\
&< \gamma m \int_{n_0+1}^{n} \frac{dt}{t\log t} + \frac{\gamma m}{(n_0+1)\log(n_0+1)}\\
&= \gamma m \log\frac{\log n}{\log(n_0+1)} + \frac{m}{10(n_0+1)\log(n_0+1)\log k}\\
&< \frac{m}{10} + \frac{m}{20(\log 2)^2}\\
&< 0.3m,
\end{aligned}$$

so that the number of remaining edges is larger than $0.7m > 7n\log n\log k$. On the other hand, the number of remaining vertices is n_0, and hence the number of remaining edges is at most n_0^k, implying $7n\log n\log k < n_0^k \leqslant n$, which leads to a contradiction.

Our claim is therefore established, and it remains to notice that when the process of dropping the vertices stops, we are left with an induced subgraph whose minimal degree satisfies the

required condition; moreover, the computation above shows that the number of edges in this subgraph is larger than $0.9m$. □

We note that the factor $\log|V'|$ in the statement of Lemma 4.1 is indeed required. To see this, for integer $N \geqslant 2$ consider the graph $G = (V, E)$ on the vertex set $V = [N]$, whose edges are pairs of vertices (i, j) with $ij \leqslant N$. It is easy to check that $|E|$ is of the order of magnitude $N \log N$. Let $G' = (V', E')$ be an induced subgraph of G. Since the largest vertex in G' is at least $|V'|$, the degree (even in G) of this vertex is at most $\frac{N}{|V'|} = O(\frac{|E|}{|V'|\log|V'|})$.

Proof of Theorem 1.3. Consider the 4-uniform hypergraph G whose vertices are lines in $\mathbb{R}^4$, parallel to the coordinate axes and passing through the points of A, and whose edges are quadruples of such lines which meet in a point of A. (Loosely speaking, the edges of G are the points of A.) Applying Lemma 4.1 we find a subset $A' \subseteq A$ with $|A'| \geqslant 0.9|A|$ such that every line, parallel to a coordinate axis and passing through a point of A', contains

$$\Omega\left(\frac{|A'|}{(|A'_1| + |A'_2| + |A'_3| + |A'_4|)\log|A'|}\right)$$

points of A'.

To simplify the notation we replace A by A' and we assume that $|A_i| \leqslant |A_4|$ for $i \in [1, 3]$; thus every line, parallel to a coordinate axis and passing through a point of A, contains $\Omega(|A|/(|A_4|\log|A|))$ points of A. Furthermore, shifting A appropriately we arrange it so that $0 \in A$ and all points of A have non-negative last coordinate.

For $x \in \mathbb{R}^4$ by $A(x)$ we denote the set of all those points from A lying on the line, parallel to L_4 and passing through x; that is, $A(x)$ is the set of those points of A, coinciding with x in the first three coordinates. In view of $0 \in A$ and since the last coordinates of all points from $A(x)$ are pairwise distinct and non-negative, if $A(x)$ is non-empty then

$$(A(x))^+ \subseteq (A - A)^+ \quad \text{and} \quad |(A(x))^+| = |A(x)| = \Omega\left(\frac{|A|}{|A_4|\log|A|}\right). \tag{4.1}$$

Choose an octant $\mathcal{O}$ of P_4 that contains at least $|A_4|/8$ points of A_4, and let $B := A_4 \cap \mathcal{O}$. If $\mathcal{O}$ is either the positive or the negative octant of P_4, then, in view of Lemma 2.1 and taking into account that $0 \in B$, we obtain

$$|(A - A)^+| \geqslant |(A_4 - A_4)^+| \geqslant |(B - B)^+| \geqslant |B|.$$

Furthermore, by Lemma 2.2 we have

$$|(A - A)^+| \geqslant \frac{1}{4^4}\frac{|A|^4}{|A_1||A_2||A_3||A_4|} = \Omega\left(\frac{|A|^4}{|B|^4}\right).$$

Combining the two last estimates we get

$$|(A - A)^+| \geqslant \Omega\big(|A|^{4/5}\big).$$

We now assume that $\mathcal{O}$ is neither the positive nor the negative octant of P_4. For each $i \in [1, 4]$ we write

$$P_i^+ := \{(x_1, \ldots, x_4) \in \mathbb{R}^4 : x_i \geqslant 0\} \quad \text{and} \quad P_i^- := \{(x_1, \ldots, x_4) \in \mathbb{R}^4 : x_i \leqslant 0\}.$$

Renumbering the coordinate axes, we can assume that either $\mathcal{O} = P_1^- \cap P_2^- \cap P_3^+ \cap P_4$, or $\mathcal{O} = P_1^+ \cap P_2^+ \cap P_3^- \cap P_4$ holds, and we now consider these two cases separately.

Case 1. Suppose first that $\mathcal{O} = P_1^+ \cap P_2^+ \cap P_3^- \cap P_4$. In keeping with our standard notation, by B_1, B_2, and B_3 we denote the orthogonal projections of B on the hyperplanes P_1, P_2, and P_3, respectively.

By Lemmas 2.2 and 2.1 we have

$$|(A-A)^+| = \Omega\left(\frac{|A|^4}{|B|^4}\right) \tag{4.2}$$

and

$$|(A-A)^+| \geqslant |(B-B)^+| = \Omega\left(\frac{|B|^3}{|B_1||B_2||B_3|}\right). \tag{4.3}$$

By Lemma 2.1 and Theorem 1.2 we have

$$|(A-A)^+| \geqslant |(B-B)^+| \geqslant |(B_1-B_1)^+| = \Omega\big(|B_1|^{2/3}\big), \tag{4.4}$$

and similarly

$$|(A-A)^+| = \Omega\big(|B_2|^{2/3}\big). \tag{4.5}$$

We notice next that, for any $b \in B$, the first and second coordinates of any point from $(A(b))^+$ are equal to those of b, and hence, if $b' \in B$ and $b'' \in B$ have different projections on P_3, then the sets $(A(b'))^+$ and $(A(b''))^+$ are disjoint. Therefore, in view of (4.1), we get

$$|(A-A)^+| = \Omega\left(|B_3|\frac{|A|}{|B|\log|A|}\right). \tag{4.6}$$

Raising (4.2)–(4.6) to the powers $1, 2, 3, 3$, and 2, respectively, and multiplying out, we obtain

$$|(A-A)^+|^{11} = \Omega\big(|A|^6/(\log|A|)^2\big). \tag{4.7}$$

Case 2. Suppose now that $\mathcal{O} = P_1^- \cap P_2^- \cap P_3^+ \cap P_4$. For $i \in [1,3]$, let l_i denote the number of points of the orthogonal projection of B on L_i. By Lemmas 2.1 and 2.2 we have

$$|(A-A)^+| = \Omega\left(\frac{|A|^4}{|B|^4}\right), \tag{4.8}$$

$$|(A-A)^+| \geqslant |(B-B)^+| = \Omega\left(\frac{|B|^3}{|B_1||B_2||B_3|}\right), \tag{4.9}$$

$$|(A-A)^+| \geqslant |(B_1-B_1)^+| = \Omega\left(\frac{|B_1|^2}{l_2 l_3}\right), \tag{4.10}$$

and similarly,

$$|(A-A)^+| = \Omega\left(\frac{|B_2|^2}{l_1 l_3}\right). \tag{4.11}$$

Evidently, we also have

$$|(A-A)^+| \geqslant l_1 \tag{4.12}$$

and

$$|(A-A)^+| \geqslant l_2, \tag{4.13}$$

and since B_3 is contained in the negative quadrant of $P_3 \cap P_4$, in view of $0 \in B_3$ we have

$$|(A-A)^+| \geqslant |(B_3-B_3)^+| \geqslant |B_3|. \tag{4.14}$$

Furthermore, notice that for any $b \in B$ the third coordinate of any point from $(A(b))^+$ is equal to that of b, and hence if $b' \in B$ and $b'' \in B$ differ in the third coordinate, then the sets $(A(b'))^+$ and $(A(b''))^+$ are disjoint. Thus, it follows from (4.1) that

$$|(A-A)^+| \geqslant \Omega\left(l_3 \frac{|A|}{|B|\log|A|}\right). \tag{4.15}$$

Squaring (4.9), (4.14), and (4.15) and multiplying out the resulting estimates, (4.8), and (4.10)–(4.13), we get (4.7). □

Acknowledgement

We are grateful to Noga Alon for communicating to us the present proof of Lemma 4.1.

References

[1] Balasubramanian, R. and Soundararajan, K. (1996) On a conjecture of R. L. Graham. *Acta Arith.* **75** 1–38.

[2] Granville, A. and Roesler, F. (1999) The set of differences of a given set. *Amer. Math. Monthly* **106** 338–344.

[3] Marica, J. and Schönheim, J. (1969) Differences of sets and a problem of Graham. *Canad. Math. Bull.* **12** 635–637.