ON COMBINATORIAL COMPLEXITY OF CONVEX SEQUENCES

A. Iosevich, S. Konyagin, M. Rudnev, and V. Ten

July 19, 2004

Abstract. We show that the equation

$$s_{i_1} + s_{i_2} + \cdots + s_{i_d} = s_{i_{d+1}} + \cdots + s_{i_{2d}}$$

has $O\left(N^{2d-2+2^{-d+1}}\right)$ solutions for any strictly convex sequence $\{s_i\}_{i=1}^N$ without any additional arithmetic assumptions. The proof is based on weighted incidence theory and an inductive procedure which allows us to effectively deal with higher dimensional interactions.

The terminology "combinatorial complexity" is borrowed from [CEGSW90] where much of our higher dimensional incidence theoretic motivation comes from.

Section 1: Introduction and statement of results

Consider a sequence of real numbers $\{s_i\}_{i=1}^N$. It is a classical problem in number theory to determine the number $\mathfrak{N}_d = \mathfrak{N}_d(N)$ of solutions of the equation

$$s_{i_1} + s_{i_2} + \cdots + s_{i_d} = s_{i_{d+1}} + \cdots + s_{i_{2d}}. \tag{1.1}$$

The number of solutions $\mathfrak{N}_d$ will certainly depend on geometric and arithmetic properties of the sequence $\{s_i\}$. A trivial example is if $s_i = i$, when the number of solutions of (1.1) is approximately N^{2d-1}. Here and throughout the paper the notations $a \lesssim b$, or $a = O(b)$ means that there exists $C > 0$ such that $a \le Cb$, and $a \approx b$ means that $a \lesssim b$ and $b \lesssim a$. Besides, $a \lessapprox b$, with respect to a large parameter N, means that for every $\epsilon > 0$ there exists $C_\epsilon > 0$ such that $a \le C_\epsilon N^\epsilon b$.

1991 *Mathematics Subject Classification.* Primary 11D45, 11L07; Secondary 52B55.

Key words and phrases. "Hard" Erdös problems, geometric complexity, incidences, convex sequences, diophantine equations, exponential sums.

Research supported in part by the NSF grant DMS02-45369, the Nuffield Foundation grant NAL/00485/A, and the EPSRC grant GR/S13682/01. The second author was supported by the INTAS grant 03-51-5070.

Typeset by $\mathcal{AMS}$-TEX

More interesting bounds are available if the sequence $\{s_i\}$ is a strictly convex in the sense that the sequence of differences $\{s_{i+1}-s_i\}$ is strictly increasing, or, equivalently, the set of points $\{(i,s_i)\}$ lies on a strictly convex curve in $\mathbb{R}^2$. For example, if $s_i=i^2$ and $d\geq 4$, one has $\mathfrak{N}_d \lesssim N^{2d-2}$. The same estimate with an appropriate power of $\log(N)$ holds in the case $d=2$ and $d=3$. This example shows that for a general strictly convex sequence, the best general upper bound for $\mathcal{N}_d$ one can hope for is $\mathcal{N}_d \lessapprox N^{2d-2}$.

Under additional arithmetic assumptions, the situation may change drastically. For example, it is known that if $s_i=i^k$ and $k>>d$, $\mathcal{N}_d \approx N^d$, and in fact the equation (1.1) only has trivial solutions. See [HB02] and the references contained therein. A non-integer example is given by $s_i=\sqrt{k_i}$, where $\{k_i\}$ is a sequence of square-free positive integers. A theorem due to Besicovitch ([Bes40]) says that these numbers are linearly independent over the field of rationals $\mathbb{Q}$. It follows that $\mathcal{N}_d \approx N^d$ in this case as well.

The examples of the previous paragraph are misleading in the sense that they lead to good estimates for $\mathcal{N}_d$ based on specific arithmetic properties of the sequence. The main thrust of this paper is to obtain the best possible upper bound on $\mathcal{N}_d$ under the assumption of strict convexity only, without any additional arithmetic or curvature hypotheses. This is achieved using geometric combinatorics.

As we indicate above, it is reasonable to conjecture that for every strictly convex sequence $\{s_i\}_{i=1}^N$, $\mathcal{N}_d \lessapprox N^{2d-2}$. We prove that this estimate is asymptotically true with an exponentially vanishing error in the exponent as d tends to infinity. More precisely, we show (see Theorem 1 below) that $\mathcal{N}_d \lesssim N^{2d-2+2^{-d+1}}$. Konyagin ([Ko00]) proved this estimate in the case $d=2$. More precisely, he showed that

$$\mathcal{N}_2 \lesssim N^{\frac{5}{2}}. \tag{1.2}$$

The equation (1.1), with $d=2$, arises if one tries to obtain a lower bound for the L_1 norm of trigonometric polynomials, see Karatsuba ([Kar98]). Namely, if $\{s_i\}_{i=1}^N$ is a convex sequence, let $\Delta(N)=\mathfrak{N}_2(N)/N^3$. Then

$$\int_0^1 \left| \sum_{j=1}^N c_j e^{2\pi \iota s_j x} \right| dx \gtrsim \Delta^{-\frac{1}{2}}(N), \tag{1.3}$$

for any array of complex unimodular coefficients c_j.

While the proof in [Ko00] is based on geometric incidence theory, Garaev ([Gar00]) developed an alternate, Fourier analytic approach to the proof of (1.2). Unification of these different points of view should lead to further progress on this problem. We hope to take up this issue in a subsequent paper.

When $d>2$, one is naturally led to consider an inductive procedure, as an alternative to higher dimensional incidence theory, where serious topological obstructions often arise. It turns out that the inductive step requires the use an appropriate weighted version of the Szemerédi-Trotter incidence theorem (see Theorem 3 below), essentially due to Szekely

([Sz97]). Unfortunately, a direct application of this weighted incidence result leads to a rather weak bound for $\mathcal{N}_d$ and an ad hoc reduction procedure is needed to replace maximal weights by average ones, resulting in a much better exponent. An effective handling the weights effectively is the key technical aspect of this paper. It requires an appropriate divide-and-conquer approach, described in Lemma 6 below.

Notation and statement of results. Fix a convex sequence $\{s_i\}_{i=1}^N$, N large. Let $B \equiv \{1, 2, \ldots, N\}$, which henceforth shall be referred to as the *base set*. Let $f : \mathbb{R} \to \mathbb{R}$ be a fixed strictly convex function such that $f(i) = s_i$. Let $S = f(B) = \{s_1, \ldots, s_N\}$.

The bounds for the quantity $\mathfrak{N}_d$ will be obtained by studying the sumset

$$dS \equiv S + \cdots + S = \{x : x = s_{i_1} + \cdots + s_{i_d},\ (i_1, \ldots, i_d) \in B^d\}. \tag{1.4}$$

Given $x \in dS$, define its multiplicity, or *weight*

$$\nu_d(x) = |\{(i_1, \ldots, i_d) \in B^d : s_{i_1} + \ldots + s_{i_d} = x\}|. \tag{1.5}$$

Here and throughout the paper the notation $|\cdot|$ stands for cardinality of a finite set. The quantity $\nu_d(x)$ will be referred to as the *weight distribution* function.

Clearly there is an L_1 relation

$$\sum_{x \in dS} \nu_d(x) = N^d, \tag{1.6}$$

the right-hand side being referred to as the *net weight*. Our goal is to obtain an L_2 bound for $\nu_d(x)$, since

$$\mathfrak{N}_d = \sum_{x \in dS} \nu_d^2(x). \tag{1.7}$$

Let $dS = \{x_1, x_2, \ldots, x_t, \ldots\}$ be ordered such that for any $x_t \in dS, \nu_d(x_t) \geq \nu_d(x_{t+1})$, if x_{t+1} is defined. It turns out that in order to estimate $\mathcal{N}_d$, it is sufficient to have a lower bound for cardinality $|dS|$ and a majorant for the weight distribution function. The former lower bound has been obtained in [ENR99] and does not require the techniques of this paper, yet it will be recovered in slightly different way and used in the framework of our proof.

Let $\mathfrak{n}_d(t) = \nu_d(x_t)$. The inverse, also a decreasing function $\mathfrak{n}_d^{-1}$ would provide the bound[1]

$$\mathfrak{n}_d^{-1}(\tau) \geq |\ dS_\tau \equiv \{x \in dS : \nu_d(x) \geq \tau\}|. \tag{1.8}$$

Our main result is the following.

[1]Note that $\mathfrak{n}_d^{-1}$ is simply the distribution function for $\mathfrak{n}_d$ in the measure-theoretical sense.

Theorem 1. *For $d \geq 2$, let $\alpha_d = 2(1-2^{-d}), \beta_d = d - \frac{4}{3}(1-2^{-d})$. Then*

$$|dS| \gtrsim N^{\alpha_d}, \tag{1.9}$$

$$\mathfrak{n}_d(t) \lesssim N^{\beta_d} t^{-1/3}, \tag{1.10}$$

$$\mathcal{N}_d \lesssim N^{2d-\alpha_d}. \tag{1.11}$$

Remark. The main estimates of this paper are (1.10) and (1.11). The estimate (1.9) on the cardinality of the sumset dS has been included in the statement for the sake of completeness and is due to Elekes et al., [ENR99], Ch. 4, where it is derived after a repeated application of the classical Szemerédi-Trotter incidence theorem. However, while the estimate (1.9) can be easily derived from the estimate (1.11), the converse is not true. We shall see that the derivation of the estimate (1.11) requires application of more sophisticated arguments involving weighted incidence with appropriately chosen weights.

In the case when the set S is a subset of integers, the estimate (1.11) enables one to bound the L_p norm of trigonometric polynomials with frequencies in S.

Corollary 2. *If $S \subset \mathbb{Z}$, let*

$$P_N(\theta) = \sum_{j=1}^{N} e^{2\pi i s_j \theta}. \tag{1.12}$$

Then

$$\|P_N\|_{2d} \equiv \left(\int_0^{2\pi} |P_N(\theta)|^{2d} d\theta \right)^{\frac{1}{2d}} \lesssim N^{1-\frac{1-2^{-d}}{d}}. \tag{1.13}$$

Remark. By expanding the square we see that (1.13) is essentially an identity when $d = 1$. When $d > 1$ observe that (1.13) is much stronger than the estimate that can be obtained by interpolating the case $d = 1$ and $d = \infty$ using Holder's inequality.

Acknowledgements. The authors wish to thank M. Garaev, N. Katz, and Z. Rudnick for a number of interesting and helpful remarks, references and suggestions.

Section 2: Incidence theorems

As we mention in the introduction, the main tool used in [ENR99] and [Ko00] is the theorem of Szemerédi and Trotter ([ST83]) bounding the number of incidences between a collection of points and straight lines in the Euclidean plane. The theorem was extended to the case of points and hyper-planes or spheres (with some natural restrictions on the arrangements) by Clarkson et al. (see [CEGSW90] and the references contained therein). Incidence theory provides a set of powerful tools for solving problems in geometric combinatorics and related areas. See also books by Pach and Agarwal ([PA95]) and Matoušek ([Ma02]) for an exhaustive description of this subject and related issues. It was observed by Székely ([Sz97]) that the geometric graph theory can deliver a short and elegant proof of the following weighted version of the Szemerédi-Trotter incidence theorem in dimension two, with the set of lines generalized to a class of curves satisfying generic intersection hypotheses.[2] From now on, we shall use the terms "lines" and "curves" interchangeably.

Theorem 2. *(*[ST83]*) Let $(\mathcal{L},\mathcal{P})$ be an arrangement*[3] *of m curves and n points in $\mathbb{R}^2$. Suppose that no more than boundedly curves pass through a pair of points of $\mathcal{P}$ and that the intersection of any two curves of $\mathcal{L}$ contains a bounded number of points of $\mathcal{P}$. Then the total number of incidences*

$$I = |\{(l,p)\in\mathcal{L}\times\mathcal{P} : p\in l\}| \lesssim (mn)^{\frac{2}{3}} + m + n. \tag{2.1}$$

Under the assumptions of Theorem 2, which we shall refer to as the *simple intersection case*, the number of incidences I for an arrangement $(\mathcal{L},\mathcal{P})$ can be expressed in terms of the counting function δ_{lp}. More precisely,

$$I = \sum_{p\in\mathcal{P}} m(p) \sum_{l\in\mathcal{L},p\in\mathcal{P}} \delta_{lp}. \tag{2.2}$$

In this formula, $m(p)$ denotes the number of curves incident to a specific point p, and $\delta_{lp}=1$ if $p\in l$, and 0 otherwise.

Now let us consider the issue of weighted incidences. In this case the numbers (m,n) in Theorem 2 will have a slightly different meaning. Given some $\mu,\nu\geq 1$ (without loss of generality suppose they are integers) let us assign to each line $l\in\mathcal{L}$ and each point $p\in\mathcal{P}$ the weights $\mu(l)\in\{1,\dots,\mu\}$ and $\nu(p)\in\{1,\dots,\nu\}$, respectively, so that

$$\sum_{l\in\mathcal{L}} \mu(l) = m, \sum_{p\in\mathcal{P}} \nu(p) = n. \tag{2.3}$$

[2]There is nothing to prevent one from generalizing the ambient space $\mathbb{R}^2$ to a general two-manifold of finite genus.

[3]By the arrangement we further mean an embedding, or drawing of the curves and points in the plane.

Let us call such a weight assignment a weight distribution with *maximum weights* (μ,ν) and *net weights* (m,n). A single pair $(l,p)\in\mathcal{L}\times\mathcal{P}$ will have the weight $w_{lp}=\mu(l)\nu(p)\delta_{lp}$. Let the number of weighted incidences be defined by

$$I\equiv\sum_{p\in\mathcal{P}}m(p)\sum_{l\in\mathcal{L},p\in\mathcal{P}}w_{lp}. \tag{2.4}$$

Now the quantity $m(p)$ counts the total weight of all the curves incident to a particular point p. Note that the cardinality of the sets $\mathcal{L}$ and $\mathcal{P}$ do not enter the weighted incidence bound (2.4) at all. We shall make use of the following weighted version of Theorem 2, which can be viewed as a variant of a weighted version of the Szemeredi-Trotter incidence theorem due to Szekely ([Sz97]).

Theorem 3. *Given a simple intersection arrangements $(\mathcal{L},\mathcal{P})$ with net weight (m,n), and a weight distributions with maximum weights (μ,ν), one has*

$$I\lesssim(\mu\nu)^{\frac{1}{3}}(mn)^{\frac{2}{3}}+\nu m+\mu n. \tag{2.5}$$

Proof of Theorem 3 follows easily from Theorem 2 by a simple weight rearrangement argument and is given at the end of the paper. Note that for the right hand side of (2.5) one has

$$(\mu\nu)^{\frac{1}{3}}(mn)^{\frac{2}{3}}+\nu m+\mu n=\mu\nu\left[\left(\frac{m}{\mu}\frac{n}{\nu}\right)^{\frac{2}{3}}+\frac{m}{\mu}+\frac{n}{\nu}\right], \tag{2.6}$$

which indicates that the maximum number of weighted incidences is achieved when there are $\frac{m}{\mu}$ lines and $\frac{n}{\nu}$ points with uniformly distributed weights, equal to μ or ν, respectively.

Observe that unless the weights are distributed uniformly, neither $|\mathcal{L}|$, nor $|\mathcal{P}|$ enter the estimate (2.5). This suggests that the estimate (2.5) needs to be properly localized to achieve sufficiently sharp estimates. However, the situation changes if extra information about the weight distributions throughout $\mathcal{L}$ or $\mathcal{P}$ becomes available. It then opens up a wide variety of possibilities for decomposition and divide-and-conquer approaches, partitioning the sets $\mathcal{L}$ or $\mathcal{P}$ into pieces such that the estimate (2.4) applied to each piece of the partition leads to sharp estimates.

The following is a heuristic sketch of the proof of Theorem 1. The proof starts out with the case $d=2$, following [ENR99] and [Ko00], based on Theorem 2, and proving (2.1). Its essence is the interpretation of the estimation of $\mathfrak{N}_2$ as an incidence problem. The case $d=2$ is followed by induction on "dimension" d. The problem of estimating $\mathfrak{N}_{d+1}$ in terms of $\mathfrak{N}_d$ can also be interpreted as an incidence problem, but a weighted one. Each point in the corresponding set $\mathcal{P}$ will have weight equal to 1. However, the set $\mathcal{L}$ will be associated with the d-dimensional problem and will carry non-trivial weights, which will be in one-to-one correspondence with the weights $\nu_d(x)$ in the sumset dS. Note that the maximum weight $\mu = \sup_{x\in dS} \nu_d(x)$ for the elements of dS is trivially N^{d-1}, or less trivially $N^{d\frac{d-1}{d+1}}$ using the classical result of Andrews ([An63]) (see also [BL98]).

Theorem 4 [Andrews]. *The number of vertices of a convex lattice polytope*[4] *in* $\mathbb{R}^d$ *of volume* V *is* $O\left(V^{\frac{d-1}{d+1}}\right)$.

Returning to the sketch of proof of Theorem 1, we shall see that if m, μ are respectively net and maximum weights for the set of lines (m will be equal to N times N^d, the latter being the net weight of dS), then the cardinality $|\mathcal{L}|$ is much greater than $m\mu^{-1}$. In other words, there is a lower bound L on $|\mathcal{L}|$, so the majority of the members of $\mathcal{L}$ will carry weights which are smaller than the maximum weight μ. This allows one to use the bound for the "average" weight $\bar{\mu} = \frac{m}{L}$ (which is much smaller than μ) in the formula (2.5). This is proved in Lemma 6 below, which is central for the proof of Theorem 1 and leads quickly to the key estimates (1.10) and (1.11).

Remark on notation. In what follows, the quantities (μ, ν), will always denote weights for the incidence problem in question, the weighted arrangement $(\mathcal{L}, \mathcal{P})$ of curves and points respectively. On the other hand the notation ν_d always refers to the weight distribution function on the sumset dS. Throughout the induction process, individual weights of curves $l \in \mathcal{L}$ are in one-to-one correspondence with weights $\nu_d(x)$, for $x \in dS$.

Section 3: Proof of Theorem 1

The proof is by induction on d, starting from the case $d = 2$. Let

$$\gamma = \{(x, f(x)) : x \in [1, N]\}, \text{ and } \gamma_B\{(i, f(i)) : i \in B\} \tag{3.1}$$

The case $d = 2$.

Lemma 5. *We have*

$$|2S| \gtrsim N^{3/2}, \tag{3.2}$$

and

$$|2S_\tau| \equiv |\{x \in 2S : \nu_2(x) \geq \tau\}| \lesssim N^3\tau^{-3}. \tag{3.3}$$

Proof. Define $2B \equiv B + B$. Consider the set of points $\mathcal{P} = B \times S + \gamma_B = 2B \times 2S$ and the set of curves $\mathcal{L} = \gamma + B \times S$. Strict convexity of the curve γ implies that the arrangement $(\mathcal{L}, \mathcal{P})$ satisfies the simple intersection condition.

[4] A lattice polytope is a polytope with vertices in the integer lattice $\mathbb{Z}^d$.

Since $|\mathcal{P}| \lesssim N^2$, the number of incidences I for this arrangement can be estimated using Theorem 2:

$$I \lesssim N^{4/3}(|\mathcal{P}|)^{2/3}. \tag{3.4}$$

On the other hand, each curve of $\mathcal{L}$ contains at least N points of $\mathcal{P}$ (that is why $\mathcal{P}$ has been taken as $2B \times S$ rather than simply $B \times S$). It follows that $I \gtrsim N^3$, and

$$N|2S| \approx |\mathcal{P}| \gtrsim N^{5/2}, \tag{3.5}$$

which implies (3.2).

Let $\mathcal{P}_\tau = \{p \in \mathcal{P} : m(p) \geq \tau\}$, where $m(p)$ is the number (coinciding in this case with the total weight) of curves of the arrangement $\mathcal{L}$ intersecting at the point p. Applying estimate (3.4) for the number of incidences for the arrangement $(\mathcal{L}, \mathcal{P}_\tau)$, with $|\mathcal{P}_\tau|$ in place of $|\mathcal{P}|$, and comparing it with the lower bound $\tau|\mathcal{P}_\tau|$, we see that $\tau|\mathcal{P}_\tau| \leq I \lesssim N^{\frac{4}{3}}|\mathcal{P}_\tau|^{\frac{2}{3}}$, which implies that $|\mathcal{P}_\tau| \lesssim \frac{N^4}{\tau^3}$, hence

$$|2S_\tau| \approx N^{-1}|\mathcal{P}_\tau| \lesssim N^3\tau^{-3}, \tag{3.6}$$

as claimed in (3.3). Note that division by N above is due to the definition of $\mathcal{P} = 2B \times 2S$, and $|2B| \approx N$, as the base set B is the set of consecutive integers.

Motivated by (3.2), let $\bar{\nu}_2 = \sqrt{N}$ be the (approximate) upper bound for average weight over $2S$ (the net weight of $2S$ is proportional to N^2). By (3.3) the weight distribution function in the (ordered) set $2S$ satisfies

$$\nu_2(x_t) \lesssim \mathfrak{n}_2(t) = Nt^{-1/3}. \tag{3.7}$$

It follows that for the set $2S_{\bar{\nu}_2}$, containing those $O(N^{3/2})$ elements of $2S$, whose weights may exceed $\bar{\nu}_2$, one has

$$\sum_{x \in 2S_{\bar{\nu}_2}} \nu_2^2(x) \lesssim N^2 \int_1^{N^{3/2}} t^{-2/3} dt \approx N^{5/2}. \tag{3.8}$$

On the other hand, for the complement $2S^c_{\bar{\nu}_2}$ of $2S_{\bar{\nu}_2}$ in $2S$, where the weight does not exceed $\bar{\nu}_2$, one has

$$\sum_{x \in 2S^c_{\bar{\nu}_2}} \nu_2^2(x) \lesssim \bar{\nu}_2 \sum_{x \in 2S} \nu_2(x) \approx N^{5/2}. \tag{3.9}$$

This proves formulas $(1.9 - 1.11)$ in the case $d = 2$.

Remark. The estimates (3.8) and (3.9) are motivated as follows. One naturally partitions the domain $2S$ in two subsets. In the first subset, containing x such that $\nu_2(x) \gtrsim \bar{\nu}_2$ (where the quantity $\bar{\nu}_2$ has been obtained as the net weight divided by the lower bound for cardinality $|2S|$) one uses the (strictly decreasing, convex) majorant $\mathfrak{n}_2(t)$ for $\nu_2(x_t)$ and gets (3.8). The sum of $\nu_2^2(x)$ over the second subset, where $\nu_2(x) \lesssim \bar{\nu}_2$ is bounded by the product of the L_1 norm of the function $\nu_2(x)$ ($\approx N^2$) and the L_∞ norm $\bar{\nu}_2 = \sqrt{N}$ for $\nu_2(x)$, restricted to the latter subset. This yields (3.9). The same idea is used in the remaining part of the proof. The most difficult point is getting a tight enough majorant $\mathfrak{n}_d(t)$ in the case $d \geq 2$.

The case $d \Rightarrow d+1$. In order to characterize the weight distribution function $\nu_{d+1}(x)$, for $x \in (d+1)S$, consider the equation

$$f(i_1) + [\,f(i_2) + \ldots + f(i_{d+1})] = x. \tag{3.10}$$

Let $u \in dS$. Extend (3.10) to the system of equations

$$\begin{cases} f(i_1) + u = x, \\ \quad i_1 \;+ j = k, \end{cases} \qquad \forall\, (i_1, j, k, u, x) \in B \times 2B \times 2B \times dS \times (d+1)S. \tag{3.11}$$

Note that $(d+1)S$ is considered as a set, rather than multi-set. The elements of the set $dS = \{u_1, u_2, \ldots, u_t, \ldots\}$ are endowed with non-increasing weights, with the weight distribution function $\nu_d(u)$, which by the induction assumption should comply with (1.9 – 1.11). Besides, the L_1 norm of $\nu_d(u)$, over dS is proportional to N^d. The L_∞ norm of $\nu_d(u)$ is $O(N^{d\frac{d-1}{d+1}})$, by the Andrews theorem (Theorem 4). By (1.10), there is a majorant

$$\nu_d(u_t) \lesssim \mathfrak{n}_d(t) \;=\; N^{\beta_d} t^{-1/3}, \tag{3.12}$$

where $\beta_d = d - \frac{4}{3}(1 - 2^{-d})$. There is also the estimate (1.9) for the minimum cardinality of dS. The latter leads us to introduce the upper bound for the average weight $\bar{\nu}_d$ in dS,

$$\bar{\nu}_d \lesssim N^{d-\alpha_d}, \tag{3.13}$$

with $\alpha_d = 2 - 2^{-d+1}$.

The number of solutions of (3.10) is not greater than the number of solutions of (3.11), divided by N. On the other hand, (3.11) can be interpreted as a weighted incidence problem. Let $\mathcal{L}$ be the set of the curves, given by the translations γ_{ju} of the curve γ defined by (3.1), by some $(j, u) \in 2B \times dS$. For such $l = \gamma_{ju} \in \mathcal{L}$, let the weight $\mu(l) = \nu_d(u)$. Define the set of points $\mathcal{P} = 2B \times (d+1)S$, with unit weights. Then the number of solutions of (3.11) is bounded by the number of weighted incidences in the arrangement $(\mathcal{L}, \mathcal{P})$. In particular, if $x \in (d+1)S = s + u$, for some $s \in S$ and $u \in dS$, then clearly

$$\nu_{d+1}(x) = \sum_{(s,u)\in(S\times dS):x=s+u} \nu_d(u). \tag{3.14}$$

The formula (3.11) applies to the case $d = 2$ as well, with $u \in S$. Hence now the problem essentially boils down to the same scheme as it was in the case $d = 2$, except that *weighted* incidences should be counted in order to verify estimates (1.10) and (1.11). Verification of (1.9) is easier. It requires only the available (through the induction assumption) lower bound $|dS| \gtrsim N^{\alpha_d}$ and the use of (2.1) and was done in [ENR99] (and in the case $d = 2$, see (3.4) and the formula that follows it). The corresponding estimate (3.12) can be also obtained using the following lemma.

Lemma 6. *Let $\bar{\nu}$ be defined as in* (3.11) *above. Assuming the estimate* (3.12) *on the weight distribution function $\nu_d(u)$ in the set dS, the number of incidences for the above defined arrangement $(\mathcal{L}, \mathcal{P})$, describing the solutions of the system* (3.11) *is given by*

$$I \lesssim \bar{\nu}_d^{1/3} N^{2(d+1)/3} (N|(d+1)S|)^{2/3}, \tag{3.15}$$

Lemma 6 shows that in order to count the weighted incidences in the arrangement $(\mathcal{L}, \mathcal{P})$ by formula (2.3), instead of the maximum weight $\mu = O(N^{d\frac{d-1}{d+1}})$ in the set $\mathcal{L}$, given by the Andrews theorem, one can set $\mu = \bar{\nu}_d$, which is considerably smaller. Note that by definition of $\mathcal{L}$), its net weight boils down to $m = N^{d+1}$. In addition, every point $p \in \mathcal{P}$ has unit weight. The proof of Lemma 6 is given in the next section. We shall now use it to complete the proof of Theorem 1.

Assuming Lemma 6, we compare the estimate (3.15) with the fact that on each curve of $\mathcal{L}$ there lies at least N points of $\mathcal{P}$. It follows that $I \geq N^{d+2}$, as net weight of $\mathcal{L}$ is proportional to N^{d+1}. Comparing the powers of N, we get[5]

$$|(d+1)S| \gtrsim N^{2-2^{-d}} = N^{\alpha_{d+1}}. \tag{3.16}$$

This leads us to define the upper bound for the average weight in $(d+1)S$:

$$\bar{\nu}_{d+1} = N^{d+1-\alpha_{d+1}}. \tag{3.17}$$

Let $\mathcal{P}_\tau = \{p \in \mathcal{P} : m(p) \geq \tau\}$, where $m(p)$ is the total weight of all the curves of the arrangement $\mathcal{L}$ intersecting at the point p, see (2.2). Clearly $\mathcal{P}_\tau = 2B \times (d+1)S_\tau$, where $(d+1)S_\tau$ is the subset of $(d+1)S$, consisting of all those elements x, whose weight $\nu_{d+1}(x)$ is not smaller than τ. In order to estimate $|(d+1)S_\tau|$, formula (2.1) cannot be used, as one has to take into account the individual weight of each curve $\gamma_{ju} \in \mathcal{L}$ (equal to $\nu_d(u)$) passing through the given point p. Instead, weighted incidences have to be dealt with, and Lemma 6 formally enables one use the average weight $\bar{\nu}_d$ instead of μ in the application of the formula (2.3).

[5]As we have mentioned earlier, one can do without Lemma 6 in order to get (3.16). Namely, if I is the number of (non-weighted in this case) incidences for the arrangement $(\mathcal{L}, \mathcal{P})$ in question, then similarly to the case $d = 2$, one has $N(N|dS|) \lesssim I \lesssim (N|(d+1)S|)^{2/3}(N|dS|)^{2/3}$, which implies bound (3.16) for $|(d+1)S|$, under the assumption $|dS| \gtrsim N^{\alpha_d}$. This was done in [ENR99].

In view of this, we proceed by comparing the lower bound $\tau N|(d+1)S_\tau|$, for the number of weighted incidences for the arrangement $(\mathcal{L},\mathcal{P}_\tau)$ with (3.16), in which $|(d+1)S_\tau|$ substitutes $|(d+1)S|$. We get $\tau N|(d+1)S_\tau| \lesssim I \lesssim \bar\nu_d^{1/3} N^{2(d+1)/3}(N|(d+1)S_\tau|)^{2/3}$. By (3.13) this yields

$$|(d+1)S_\tau| \lesssim N^{1-\alpha_d}\left(\frac{N^d}{\tau}\right)^3. \tag{3.18}$$

If $\tau=\bar\nu_{d+1}$, defined by (3.17), it follows that

$$|(d+1)S_{\bar\nu_{d+1}}| \lesssim N^{\alpha_{d+1}}, \tag{3.19}$$

which is the same as the right-hand side in (3.16), and complies with (1.9). Inversion of (3.14) yields:

$$\nu_{d+1}(x_t) \lesssim \mathfrak{n}_{d+1}(t) = N^{d-\frac{1-2^{-d+1}}{3}} t^{-1/3} = N^{\beta_{d+1}} t^{-1/3}, \tag{3.20}$$

as is claimed by (1.10).

The final step of the proof follows the remark at the end of the $d=2$ section. More precisely, one partitions

$$(d+1)S = (d+1)S_{\bar\nu_{d+1}} \cup (d+1)S^c_{\bar\nu_{d+1}}, \tag{3.21}$$

into "heavy" and "light" elements, and obtains the estimate

$$\sum_{x\in(d+1)S^c_{\bar\nu_{d+1}}} \nu^2_{d+1}(x) \lesssim N^{d+1}\bar\nu_{d+1} = N^{2(d+1)-\alpha_{d+1}}, \tag{3.22}$$

along with

$$\sum_{x\in(d+1)S_{\bar\nu_{d+1}}} \nu^2_{d+1}(x) \lesssim N^{2\beta_{d+1}} \int_1^{N^{\alpha_{d+1}}} t^{-2/3}dt \approx N^{2(d+1)-\alpha_{d+1}}. \tag{3.23}$$

Estimates (3.22) and (3.23) are consistent with (1.11). Thus the proof of Theorem 1 is complete up to the verification of Lemma 6.

Section 4: Proofs of Lemma 6 and Theorem 3

Proof of Lemma 6. The objective is to partition the set

$$dS = \bigcup_{i=0}^{M} dS_i \tag{4.1}$$

into M (a fairly large number of) pieces, trying to make each one of them as large as possible, yet having control over the number of weighted incidences it can possibly be responsible for. We aim to get a bound

$$\nu_d(x) \lesssim b_i, \forall x \in dS_i, \tag{4.2}$$

for some geometrically decreasing sequence b_i approaching the quantity $\bar{\nu}_d$, defined by (3.17) and appearing in main estimate (3.15). The sequence b_i will start out from

$$b_0 = N^{d\frac{d-1}{d+1}}, \tag{4.3}$$

(the L_∞ norm of ν_d, given by the Andrews theorem[6]). The number M in (4.1) is chosen in such a way that b_M is close enough to $\bar{\nu}_d$, so that the effect of the difference between them can be swallowed by a constant in the $\lesssim$ symbol. The sequence $\{dS_i\}$ will be constructed, using the weight distribution majorant (3.12).

By general estimate (2.3) of Theorem 3 in order to prove the lemma, it suffices to show that

$$\left(\tilde{I} \equiv \sum_{i=0}^{M} b_i^{\frac{1}{3}} m_i^{2/3}\right) \lesssim \left(\bar{I} \equiv \bar{\nu}_d^{1/3} m^{2/3}\right), \tag{4.4}$$

where $m = N^d$ is the net weight of dS, and m_i is the net weight of each subset dS_i, for $i = 0, \ldots, M$. The difference between (4.4) and (3.15) is that we have dropped those powers of N in the latter estimate, which arise from net weight of $\mathcal{L}$ as well as the fact that $\mathcal{P} = 2B \times (d+1)S$ (i.e that to every $x \in (d+1)S$ there correspond at least N solutions of (3.11)). Each $dS_i \subset dS$ corresponds to the subset $\mathcal{L}_i = 2B \times dS_i$ of $\mathcal{L}$. Throughout the proof of Lemma 6, m_i would stand for net weights of dS_i only, rather than $\mathcal{L}_i$.

It is easy to verify that the linear terms coming from bound (2.3) are irrelevant. Indeed, the first linear term is $O(N^{d+1})$, being the total weight of the set of lines $\mathcal{L} = 2B \times dS$. The second linear term can be bounded via $b_i N^{d+1}$. By construction, both linear terms will be dominated by the incidence bound, reflected by the quantity $\tilde{I}$, defined by (4.4). See (4.12) at the end of the proof.

Net weights m_i of dS_i are to be estimated via b_i, using the inverse formula for the majorant (3.9), i.e

$$|\{x \in dS : \nu_d(x) \geq \tau\}| \lesssim \mathfrak{n}_d^{-1}(\tau) N^{3\beta_d} \tau^{-3}, \quad \beta_d = d - \frac{4}{3}(1 - 2^{-d}). \tag{4.5}$$

Note that the majorant (3.12) is good for nothing as far as the elements $x \in dS$, such that $\nu_d(x) \lesssim \bar{\nu}_d$ are concerned. Indeed, a calculation yields

$$\int_{\bar{\nu}_d}^{\infty} \mathfrak{n}_d^{-1}(\tau) ds \approx m, \tag{4.6}$$

[6] In fact, one can see from the proof that the use of the Andrews theorem is unnecessary: one can simply start out with $b_0 = N^d$, which is the net weight of dS.

where $m \approx N^d$ is the net weight of dS.

Also for the terms in the sum in the right-hand side of (4.4) denote

$$\tilde{I}_i \equiv b_i^{\frac{1}{3}} m_i^{2/3}. \tag{4.7}$$

The sets dS_i and the number M are to be chosen such that

$$\tilde{I}_i \lesssim N^{-\varepsilon_i} \bar{I}, \tag{4.8}$$

for some geometrically vanishing sequence of small positive numbers $\{\varepsilon_i\}_{i=0}^{M-1}$. This prompts the choice

$$M \approx \log \log N, \tag{4.9}$$

as then $\varepsilon_{M-1} \approx \varepsilon_0\, e^{-\log\log N} \approx \frac{1}{\log N}$, so for a sufficiently small, yet $O(1)$ value of ε_0,

$$N^{\varepsilon_{M-1}} \approx 1 \text{ and } \sum_{i=0}^{M-1} N^{-\varepsilon_i} \approx \int_1^{\log\log N} N^{-\varepsilon_0\, exp(-t)} dt \lesssim \int_1^\infty \frac{e^{-z}}{z} dz \approx 1. \tag{4.10}$$

Let us describe the first step of the construction. Let a number δ_0 be defined via $b_0 = N^{\delta_0}\bar{\nu}_d$. Define the weight m_0 of the set dS_0 implicitly via (4.5):

$$b_0^{1/3} m_0^{2/3} \approx N^{-\varepsilon_0} \bar{\nu}_d^{1/3} m^{2/3}, \tag{4.11}$$

which yields

$$m_0 = N^{-\frac{1}{2}(3\varepsilon_0 + \delta_0)} m. \tag{4.12}$$

Then the weight of any element x in the complement dS_0^c of dS_0 in dS should be bounded from above by some quantity b_1, which can be defined implicitly from

$$\int_{b_1}^\infty \mathfrak{n}_d^{-1}(\tau)\, ds \;=\; m_0. \tag{4.13}$$

This yields

$$b_1 = \bar{\nu}_d N^{\delta_1}, \delta_1 \frac{1}{4}(3\varepsilon_0 + \delta_0). \tag{4.14}$$

Clearly, for ε_0 small enough, say $\varepsilon_0 = \frac{1}{9}\delta_0$, one has $\delta_1 \leq \frac{1}{3}\delta_0$.

The procedure is now repeated for the set dS_0^c, where the maximum weight is bounded in terms of b_1, rather than b_0, which will result in some set dS_1 having been pulled out of it, such that the maximum weight in the complement of dS_1 in dS_0^c is bounded in terms of some b_2 (which is much smaller than b_1), and so on. After having done it $M-1$ times, the set dS will be partitioned, according to (4.1), where the last member of the partition dS_M is the complement of the union $\bigcup_{i=0}^{M-1} dS_i$ in dS. For $i=1,\ldots,M$ the maximum individual element weight in dS_i is bounded similarly to (4.14), namely

$$b_i = \bar{\nu}_d N^{\delta_i}, \;\; \delta_i \frac{1}{4}(3\varepsilon_{i-1} + \delta_{i-1}). \tag{4.15}$$

Thus, if the quantities ε_i vanish geometrically, with the ratio exceeding say 9, we have $\delta_i \leq \delta_0 e^{-i}$, $i=1,\ldots,M$.

By construction, each set of lines $\mathcal{L}_i = 2B \times dS_i$, for $i=0,\ldots,M-1$ would create the number of weighted incidences I_i for the arrangement $(\mathcal{L},\mathcal{P})$, bounded as follows:

$$I_i \lesssim N^{2d+2^{-d}-\varepsilon_i+1}. \tag{4.16}$$

See (3.15) and (4.7). Note that in comparison with (1.11) one has $d \to d+1$, which accounts for an extra N here, as the quantity $\mathfrak{N}_d$ equals N^{-1} times the number of incidences for the arrangement $(\mathcal{L},\mathcal{P})$, introduced in accordance with the system of equations (3.11), rather than equation (3.10).

As each $\varepsilon_i \leq 1$, the right hand side of the last expression will exceed the maximum for the linear term in estimate (2.3), applied to the arrangement $(\mathcal{L},\mathcal{P})$, as the latter can be bounded simply via

$$b_0 N^{d+1} \lesssim N^{\frac{2d^2}{d+1}}. \tag{4.17}$$

Finally, by (4.7)

$$b_M \lesssim \bar{\nu}_d, \tag{4.18}$$

and thus the remaining set dS_M, as well as (also by (4.10)) the union $\bigcup_{i-1}^{M-1} dS_i$ will not be responsible for more incidences than specified by the right-hand side of (3.11). This completes the proof of Lemma 6.

Proof of Theorem 3. Without loss of generality, one can assume that all the weights are integers, the net line weight m is a multiple of the maximum line weight μ, and the net point weight n is a multiple of the maximum point weight ν. Then bound (2.5) is equivalent to the bound (2.1) for the number of incidences between m/μ lines and n/ν points, provided that in the latter bound, each incidence has been counted $\mu\nu$ times. In other words, for the uniform weight distribution there is nothing to prove.

Otherwise, consider some arrangement $(\mathcal{L},\mathcal{P})$ and suppose, that the weight distribution over, say $\mathcal{P}$ is not uniform. Then there exist $p_1, p_2 \in \mathcal{P}$, such that $\nu(p_1) < \nu(p_2) < \nu$. For $p \in \mathcal{P}$ let

$$m(p) = \sum_{l\in\mathcal{L}} \mu(l)\delta_{lp}, \tag{4.19}$$

be the total weight of all the lines incident to p. If $m(p_1) > m(p_2)$, first change the weight distribution by swapping the values $\nu(p_1)$ and $\nu(p_2)$ over the points p_1 and p_2. Then modify the weight distribution by changing $\nu(p_1) \to \nu(p_1) - 1$ and $\nu(p_2) \to \nu(p_2) + 1$. If $\nu(p_1)$ has become zero, remove p_1 from $\mathcal{P}$. As the result, the weight distribution has been modified, so that the number of weighted incidences has increased, yet the net weight has stayed constant. Continue this (greedy) procedure, until the weight distribution over $\mathcal{P}$ has become uniform; then do the same thing with the set $\mathcal{L}$. At each single step, the number of incidences will have increased. However, as the result, one still ends up with bound (2.4), as only m/μ lines and n/ν points will eventually remain. This completes the proof of Theorem 3.

Section 5: Theorem 1 and inequalities for elements of special matrices

In this section we will present another approach to the proof of Theorem 1 based on construction and study of special matrices. The same idea was used in ([Ko02]) to get estimates for exponential sums over subgroups of multiplicative groups in finite fields. The proofs are similar to those in ([Ko02]), so we only sketch the arguments.

First we observe that it is enough to prove Theorem 1 in the case when each each s_i is an integer. Indeed, let $\{s_i\}_{i=1}^N$ be an arbitrary convex sequence. By the pigeon-hole principle, there are integers S_i $(i = 1, \dots, N)$ and a positive integer M, such that for $i = 1, \dots, N$ we have

$$|Ms_i - S_j| < 1/(2d). \tag{5.1}$$

Then the equality

$$s_{i_1} + s_{i_2} + \cdots + s_{i_d} = s_{i_{d+1}} + \cdots + s_{i_{2d}} \tag{5.2}$$

implies

$$S_{i_1} + S_{i_2} + \cdots + S_{i_d} = S_{i_{d+1}} + \cdots + S_{i_{2d}}. \tag{5.3}$$

Therefore, the number of solutions to equation (5.2) does not exceed the number of solutions to equation (5.3). Moreover, M can be chosen so large that $s_{i+1} - 2s_i + s_{i-1} > 1/M$ for $i = 2, \dots, N-1$. Hence, the sequence $\{S_i\}_{i=1}^N$ is also strictly convex. We see that (1.11) for integral strictly convex sequences implies its validity for all strictly convex sequences. Similarly this can be shown for inequalities (1.9) and (1.10).

So, we will assume that a sequence $\{s_i\}_{i=1}^N$ is integral and strictly convex. Fix d and take a large positive integer p. Then the equation $s_{i_1} + s_{i_2} + \cdots + s_{i_d} = x$ is equivalent to the congruence $s_{i_1} + s_{i_2} + \cdots + s_{i_d} \equiv x (\mathrm{mod} p)$. We arrange the square matrix A of order p setting $a_{k,l} = 1$ if $l - k \equiv s_i (\mathrm{mod} p)$ for some i and $a_{k,l} = 0$ otherwise.

By $a_{k,l}^{(d)}$ we denote the elements of the matrix A^d. Clearly, $\nu_d(x) = 0$ if $|x| > d\overline{s}$ where $\overline{s} = \max_i |s_i|$. It is easy to check that $a_{k,l}^{(d)} = \nu_d(l-k)$ for $|l-k| \le d\overline{s}$ provided that p is large enough. By $\{\overline{a}_1^{(d)}, \ldots, \overline{a}_p^{(d)}\}$ we denote the non-increasing rearrangement of a row of the matrix A^d (observe that it does not depend on the choice of a row because any row of A^d is a cyclic translation of any other row). Inequality (1.10) means that

$$a_t^{(d)} \lesssim N^{\beta_d} t^{-1/3}. \tag{5.4}$$

Also, for any k

$$\mathcal{N}_d = \sum_{l=1}^{p} \left(a_{k,l}^{(d)}\right)^2 = \sum_{t=1}^{p} \left(a_t^{(d)}\right)^2, \tag{5.5}$$

and (1.11) is equivalent to

$$\sum_{t=1}^{p} \left(a_t^{(d)}\right)^2 \lesssim N^{2d-\alpha_d}. \tag{5.6}$$

It is easy to see that the following equalities hold:

$$\forall k \quad \sum_{l} a_{k,l} = N, \tag{5.7}$$

$$\forall l \quad \sum_{k} a_{k,l} = N. \tag{5.8}$$

Let U be the column of size p all whose elements are equal to 1. Equality (5.7) is equivalent to $AU = NU$. This implies $A^d U = N^d U$, or

$$\forall k \quad \sum_{l} a_{k,l}^{(d)} = N^d. \tag{5.9}$$

In turn, (5.7) can be rewritten as

$$\sum_{t} a_t^{(d)} = N^d \tag{5.10}$$

followed by

$$a_t^{(d)} \leq N^d / t. \tag{5.11}$$

Estimates (5.4) and (5.11) easily imply (5.6).

To prove (5.4) we need some other properties of the matrix A which can be deduced from the Szemerédi-Trotter incidence theorem.

Lemma 7. *For any sets $K \subset \{1, 2, \dots, p\}$ and $L \subset \{1, 2, \dots, p\}$ we have*

$$\sum_{k\in K}\sum_{l\in L} a_{k,l} \lesssim N^{\frac{1}{3}}(|K|\cdot|L|)^{\frac{2}{3}} + |K| + |L|. \tag{5.12}$$

Proof. It will be more convenient to work in $\mathbb{Z}$ rather than in $\mathbb{Z}(\mathrm{mod} p)$. We note that

$$\sum_{k\in K}\sum_{l\in L} a_{k,l} \le S(K, L') \tag{5.13}$$

where $L' = L \cup (L-p) \cap (L+p)$ and $S(K, L')$ is the number of solutions to the equation

$$l - k = s_i \quad k \in K, l \in L', i \in \{1, \dots, N\}. \tag{5.14}$$

Thus, we have to show that

$$S(K, L') \lesssim N^{\frac{1}{3}}(|K|\cdot|L|)^{\frac{2}{3}} + |K| + |L|. \tag{5.15}$$

Following the proof of Lemma 5, we consider the set of points $\mathcal{P} = 2B \times L'$ and the set of curves $\mathcal{L} = \gamma + 2B \times K$. Let I be the number of incidences for this arrangement. We have

$$|\mathcal{P}| \le 2N|L'|, \quad |\mathcal{L}| \le N|K|, \quad I = NS(K, L'). \tag{5.16}$$

Using the Szemerédi-Trotter incidence theorem, we get (5.15). Combining (5.13) and (5.15), we complete the proof of Lemma 7.

The crucial estimate (5.4) can be deduced from Lemma 7 and (5.7)-(5.9) by induction on d similarly to the proof of Lemma 19 in ([Ko02]).

References

[An63] G. E. Andrews, *A lower bound for the volume of strictly convex bodies with many boundary lattice points*, Trans. Amer. Math. Soc. **106** (1963), 270–279.

[Bes40] A. S. Besicovitch, *On the linear independence of fractional powers of integers*, J. London Math. Soc. **15** (1940), 3–6.

[BL98] I. Bàràny and D. G. Larman, *The convex hull of the integer points in a large ball.*, Math. Ann. **312** (1998), 167–181.

[CEGSW90] K. Clarkson, H. Edelsbrunner, L. Guibas, M. Sharir, and E. Welzl, *Combinatorial complexity bounds for arrangements of curves and surfaces*, Discrete and Computational Geometry **5** (1990), 99–160.

[ENR99] G. Elekes, M. Nathanson, and I. Ruzsa, *Convexity and sumsets*, Journal of Number Theory **83** (1999), 194–201.

[Gar00] M. Garaev, *On lower bounds for the L^1-norm of exponential sums*, Mathematical Notes **68** (2000), 713–720.

[HB02] D. R. Heath-Brown, *Counting rational points on algebraic varieties*, (preprint) (2002).

[Kar98] A. A. Karatsuba, *An estimate of the L^1-norm of an exponential sum*, Mat. Zametki **64** (1998), 465–468.

[Ko00] S.V. Konyagin, *An estimate of the L_1-norm of an exponential sum (in Russian)*, in The Theory of Approximations of Functions and Operators. Abstracts of Papers of the International Conference, Dedicated to Stechkin's 80th Anniversary, Ekaterinbourg (2000), 88–89.

[Ko02] S. Konyagin, *Estimates for trigonometric sums over subgroups and for Gaussian sums (in Russian)*, in IV International conference "Modern problems in Number theory and its applications", Topical problems, part III. Moscow (2002), 86–114.

[Ma02] J. Matoušek, *Lectures on Discrete Geometry*, Springer-Verlag, New York, 2002.

[PA95] J. Pach and P. Agarwal, *Combinatorial Geometry*, Wiley-Interscience Series in Discrete Mathematics and Optimization, 1995.

[Sz97] L. Székely, *Crossing numbers and hard Erdös problems in discrete geometry*, Combinatorics, Probability, and Computing **6** (1997), 353–358.

[ST83] E. Szemerédi and W. Trotter, *Extremal problems in discrete geometry*, Combinatorica **3** (1983), 381–392.