INCIDENCES AND EXTREMAL PROBLEMS ON FINITE POINT SETS

BY BENJAMIN LUND

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Computer Science

Written under the direction of

Shubhangi Saraf

and approved by

New Brunswick, New Jersey

May, 2017

ABSTRACT OF THE DISSERTATION

Incidences and extremal problems on finite point sets

by Benjamin Lund

Dissertation Director: Shubhangi Saraf

This thesis consists of three chapters, each addressing a different collection of problems on the extremal combinatorics of finite point sets.

The first collection of results is on the number of flats of each dimensions spanned by a set of points in $\mathbb{R}^d$. These results generalize a theorem of Beck [7] from 1983, and answer a question of Purdy [28] from 1995. We also apply the ideas behind the main results of the chapter to generalize an incidence bound between points and planes proved by Elekes and Tóth [23] to all dimensions. With the exception of the generalization of the Elekes-Tóth incidence bound, all of the material in this chapter has previously appeared as [43].

The second collection of results is on the set of perpendicular bisectors determined by a set of points in the plane. We show that if P is a set of points in $\mathbb{R}^2$ such that no line or circle contains more than a large constant fraction of the points of P, the the pairs of points of P determine a substantially superlinear number of distinct perpendicular bisectors. This is the first substantial progress toward a conjecture of the author, Sheffer, and de Zeeuw [46] that such a set of points must determine $\Omega(n^2)$ distinct perpendicular bisectors. This chapter also includes a new proof of a known result on an old question Erdős [25] on the distances between pairs of points in the plane. This chapter is [44].

The third collection of results concerns the set of flats spanned by a set of points in $\mathbb{F}_q^d$. For a set of points P in $\mathbb{F}_q^2$, this result implies that, for any $\varepsilon > 0$, if $|P| > (1+\varepsilon)q$, then $\Omega(q^2)$ lines each contain at least two points of P. We obtain a tight generalization of this statement to all dimensions, as well as a more general result for block designs. We use this theorem to improve a result of Iosevich, Rudnev, and Zhai [39] on the distinct areas of triangles determined by points in $\mathbb{F}_q^2$. This chapter is joint work with Shubhangi Saraf, and has been published as [45].

Acknowledgements

This thesis includes joint work with Shubhangi Saraf. The bulk of Chapter 2 will appear in Combinatorica, and is available online as [43]. Chapter 3 is [44], and has been submitted for publication. Chapter 4 is joint work with Shubhangi Saraf, published as [45] in SIAM Journal on Discrete Mathematics.

Many thanks to all of my collaborators, including Imre Bárány, Abdul Basit, Julien Bureaux, Javier Cilleruelo, Davey Fitzpatrick, Brandon Hanson, Alex Iosevich, Alexander Magazinov, Thang Pham, George Purdy, Oliver Roche-Newton, Misha Rudnev, Shubhangi Saraf, Adam Sheffer, Justin Smith, Le Anh Vinh, Charles Wolf, and Frank de Zeeuw.

Thanks to George Purdy for introducing me to discrete geometry when I was a student at University of Cincinnati. Thanks to Imre Bárány for hosting me at Rényi Institute, and for teaching me new problems and techniques in convex geometry. Thanks most of all to my advisor Shubhangi Saraf for her patience, guidance, and encouragement.

Table of Contents

Chapter 1

Introduction

This thesis is about the extremal combinatorics of finite point sets. Two classical examples of questions of this type are: How few lines can pass through pairs of points in a set of n points in the real affine plane, no more than k of which are collinear? How few distances can occur between pairs of points in a set of n points in the Euclidean plane?

The first of these questions was asked by Erdős [26], who conjectured that such a set of n points in $\mathbb{R}^2$, no more than k of which lie on any single line, must span $\Omega(n(n-k))$ lines in total. This conjecture was proved by Beck in 1982 [7].

The second question is the celebrated distinct distance problem of Erdős [25]. It is a classical result in number theory that the count of natural numbers less than n that are the sum of squares is $\Theta(n/\sqrt{\log n})$; hence, a $\sqrt{n} \times \sqrt{n}$ section of the integer lattice determines $\Theta(n/\sqrt{\log n})$ distinct distances. In 1946, Erdős conjectured that any set of n points in $\mathbb{R}^2$ must determine $\Omega(n/\sqrt{\log n})$ distinct distances. This *distinct distances* problem remained wide open until 2010, when Guth and Katz [35] showed that a set of n points in $\mathbb{R}^2$ must determine $\Omega(n/\log n)$ distinct distances.

A key notion for these problems is that of *incidence*: we say that a point p is incident to a geometric object Λ if $p \in \Lambda$, and we are frequently interested in placing an upper bound on the number of incidences that can occur between a set of points and a set of geometric objects from some fixed family. In the simplest case, we consider incidences between points and lines. The breakthroughs of Beck and of Guth and Katz both relied on progress on point-line incidence bounds.

Beck's theorem can be proved using the Szemerédi-Trotter theorem [59].[1] Let P be

[1] In fact, Beck proved a similar, but weaker, incidence bound to prove his theorem. The papers of

a set of n points in $\mathbb{R}^2$; we say that a line ℓ is r-rich if it contains at least r points of P. The Szemerédi-Trotter theorem states that the number of r rich lines is bounded above by $O(n^2/r^3 + n/r)$; dually, this implies the same bound for a set of r-rich points determined by a set of lines. It is possible to see that this theorem cannot be improved, by taking the point set to be a $\sqrt{n} \times \sqrt{n}$ section of the integer lattice. The Szemerédi-Trotter theorem has been very useful in combinatorial geometry, additive combinatorics, and computational geometry; for examples, see [50, section 4].

In their approach to the distinct distances problem, Guth and Katz proved a stronger bound than the Szemerédi-Trotter theorem, under the additional assumption that not too many lines lie in any single plane or other doubly ruled surface. In particular, they showed that, if L is a set of n lines in $\mathbb{R}^3$ of which no more than $\sqrt{n}$ lie in any single plane or regulus, then the number of points that lie in at least $r < \sqrt{n}$ lines is bounded above by $O(n^{3/2}/r^2)$. In proving this theorem, Guth and Katz introduced a new partitioning technique for real space based on polynomials, which has had a huge recent impact on the field – for example, polynomial partitioning is essential to the results in Chapter 3 of this thesis.

This thesis is organized into three main chapters, each dealing with a different collection of questions about points. Chapter 2 solves long standing problems relating to higher dimensional versions of Beck's theorem. Chapter 3 investigates point sets that determine few perpendicular bisectors, and relates this to a question of Erdős on the distances determined by a point set. Chapter 4 considers finite field variants of Beck's theorem, which follows the pioneering work of Bourgain, Katz, and Tao [11] in considering finite field analogues to classical problems in combinatorial geometry. The remainder of the introduction gives additional background for each of these Chapters.

1.1 The flats spanned by a set of points in real space

Let P be an arbitrary set of n points in $\mathbb{R}^d$.

In 1982, Beck [7] proved the following theorem, first conjectured by Erdős [26].

Szemerédi and Trotter and of Beck appeared in the same issue of Combinatorica.

Theorem 1 (Beck-Erdős)**.** *If no line contains more than k points of P, then the number of lines that each contain at least two points of P is $\Omega(n(n-k))$.*

Here, the condition that a line ℓ contains at least two points of P is equivalent to the condition that ℓ is *spanned* by two points of P; that is, there are points $x, y \in P$ such that each point $z \in \ell$ is an affine combination of x, y. We may also consider the set of higher-dimensional flats spanned by P; we say that a k-flat Γ is spanned by P if it contains $k+1$ affinely independent points of P. For example, a plane is spanned by P if it contains 3 non-collinear points of P.

One important implication of Theorem 1 is that, if no line contains more than any fixed, constant fraction of the points of P, then P spans $\Omega(n^2)$ lines; that is, a positive fraction of the maximum possible. To obtain a similar result for planes, it is not sufficient to suppose that no more than an arbitrary constant fraction of the points are contained in any plane, as shown by the following example. Suppose that P is contained in the union of two skew (non-intersecting) lines in $\mathbb{R}^3$. Since any plane that contains two points on one of the lines must contain the entire line, it is not hard to see that the number of planes spanned by P is n. On the other hand, no plane contains more than $n/2+1$ points of P.

One unusual characteristic of a set of points contained in the union of two skew lines is that it spans many more lines than planes. For each $1 \leq k \leq d-1$, denote by f_k the number of k-flats spanned by P. In 1986, Purdy [52] proved that either P is contained in a plane, or the union of two lines, or $f_2 = \Omega(f_1)$. He also conjectured specific conditions under which the exact inequality $f_2 \geq f_1$ should hold, and more generally asked for conditions under which $f_k \geq f_{k-1}$. A survey article by Erdős and Purdy [28] contains some additional details of Purdy's conjectures and results on the problem.

The main result of Chapter 2 is a nearly complete answer to Purdy's question. For a finite set of points Q, denote by $K(Q)$ the least t such that Q is contained in the union of a set of flats, each of dimension at least 1, whose dimensions add up to t. Denote by g_i the size of the largest subset $Q \subseteq P$ such that $K(Q) \leq i$, for each $0 \leq i \leq d-1$.

Theorem 2. *For each $k \geq K(P)$ such that $f_{k-1} > 0$, we have $f_{k-1} > f_k$. For each $2 \leq k < K(P)$, there is a constant c_k depending only on k such that either $g_{k-1} > n - c_k$, or $f_{k-1} > f_k$.*

For example, either all but c_2 points of P are contained in a plane, all but c_2 points of P are contained in the union of two lines, or P spans more planes than lines. Note that this is stronger than the conclusion of Purdy's theorem mentioned above.

Theorem 2 answers Purdy's question, except in the case in which all but fewer than c_k points of P are contained in a union of flats whose dimensions sum to k. Chapter 2 includes examples showing that it is not possible to get a more precise answer using only the hypotheses of Theorem 2. These examples give lower bounds on the values of c_k for which Theorem 6 can hold; in particular, we show that c_k must depend at least linearly on k, and conjecture that c_k must have an exponential dependence on k.

Key to proving Theorem 2 are the following matching upper and lower bounds on the number of k-flats spanned by P.

Theorem 3. *For $k < K(P)$,*

$$f_k = \Theta\left(\prod_{i=0}^{k}(n - g_i)\right).$$

The case $k = 1$ of Theorem 3 is exactly Theorem 1, and hence Theorem 3 is a very strong generalization of Theorem 1.

As a further application of the idea of essential dimension, Chapter 2 generalizes a bound on the number of incidences between 2-flats and points proved by Elekes and Tóth [23] to bound incidences between k-flats and points.

The proofs in Chapter 2 rely on elementary projective geometry, combinatorics, and the Szemerédi-Trotter bound on the number of incidences between points and lines in the real plane.

1.2 Perpendicular bisectors and distinct distances

While Chapter 2 deals with affine properties of point sets, Chapter 3 considers properties induced by the Euclidean distance, here denoted $\|\cdot\|$.

A pair of distinct points $a, b \in \mathbb{R}^2$ determines the perpendicular bisector

$$\mathcal{B}(a, b) = \{x \in \mathbb{R}^2 : \|x - a\| = \|x - b\|\}.$$

In Chapter 3, we give bounds on the number of perpendicular bisectors determined by an arbitrary set of points in the Euclidean plane, together with an application to an old problem of Erdős on distances determined by a set of points in the plane.

Let P be a set of n points in $\mathbb{R}^2$. Denote by

$$\mathcal{B}(P) = \{B(a, b) : a, b \in P, a \neq b\}$$

the set of bisectors determined by P. In the spirit of the Erdős distinct distances problem, it is natural to ask how few lines can occur as the perpendicular bisectors of a set of n points in the plane. Since the vertices of a regular n-gon determine exactly n bisectors, and each point determines $n - 1$ bisectors with the remaining points, it is easy to give a nearly complete answer to the naive version of this question.

However, it seems that the only way to have $o(n^2)$ bisectors is by placing nearly all of the points on a single line or circle, and the purpose of Chapter 3 is to make progress on the following conjecture.

Conjecture 4. *Either $n/2$ points of P lie on a single line or circle, or $|\mathcal{B}(P)| = \Omega(n^2)$.*

There is nothing special about the constant $1/2$ in this conjecture; it may be replaced by anything less than 1. This conjecture is reminiscent of the situation for lines spanned by a point set; just as Beck proved that either nearly all of the points of P are collinear, or the number of lines spanned is quadratic, we hope to prove the same for perpendicular bisectors.

We approach the problem of finding an upper bound on the number of bisectors determined by a point set by placing a lower bound on the *bisector energy*, defined by

$$|\mathcal{Q}| = \{(a, b, c, d) \in P^4 : a \neq b, c \neq d, \mathcal{B}(a, b) = \mathcal{B}(c, d)\}.$$

This is analogous to the additive energy, which is of great importance in additive combinatorics. Given an upper bound on $|\mathcal{Q}|$, a standard argument using the Cauchy-Schwarz inequality implies a corresponding lower bound on $|\mathcal{B}(P)|$.

As with $|\mathcal{B}(P)|$, the easy bound of $O(n^3)$ on $|\mathcal{Q}|$ is met when P is the set of vertices of a regular n-gon. Hence, rather than bounding the full energy $|\mathcal{Q}|$, we bound the energy

$$|\mathcal{Q}^*| = \{((a, b), (c, d)) \in \Pi \times \Pi : \mathcal{B}(a, b) = \mathcal{B}(c, d)\}$$

of a carefully chosen subset $\Pi \subseteq P \times P$ of the pairs of points of P.

Using this refined energy bound, we make some progress toward Conjecture 4.

Theorem 5. *For any $\delta, \varepsilon > 0$, either a single circle or line contains $(1 - \delta)n$ points of P, or*

$$|\mathcal{B}(P)| = \Omega(n^{52/35-\varepsilon}).$$

Chapter 3 also includes applications of ideas used in the proof of Theorem 5 to a classical question posed by Erdős on the distances determined by a set of points in the plane.

The famous *distinct distances* conjecture of Erdős that P determines at least $\Omega(n/\sqrt{\log n})$ distinct distances. He also made the stronger *pinned distances* conjecture that there is always a point $x \in P$ such that the points of $P \setminus \{x\}$ determine at least $\Omega(n/\sqrt{\log n})$ distinct distances from x. While Guth and Katz [35] closed the gap for the distinct distances problem to $\sqrt{\log n}$, the best result on the pinned distances problem is by Katz and Tardos [41], who proved that there is always a point from which there are at least $\Omega(n^{0.864})$ distinct distances. Chapter 3 contains a completely new proof of a slightly weaker result on the pinned distance problem.

The proofs Chapter 3 rely on a new upper bound on the number of incidences between points and certain bounded degree real algebraic varieties. The proof of these incidence bounds rely on the polynomial partitioning technique developed by Guth and Katz [35] in their result on the Erdős distinct distance problem, along with further development of this approach by Fox et. al. [29].

1.3 Flats spanned by a set of points in finite space

In Chapter 4, we prove incidence bounds and Beck-type theorems for point sets in vector spaces over finite fields of odd characteristic. We will work in $\mathbb{F}_q^d$, the d dimensional vector space over the field with q elements, for an arbitrary prime power q.

It is easy to see that the Szemerédi-Trotter theorem and Beck's theorems do not hold in this setting. Suppose that P consists of all of the points in $\mathbb{F}_q^2$. Certainly, $|P| = q^2$. On the other hand, the total number of lines spanned by P is $q^2 + q$, and each of these lines contains q points of P. Beck's theorem, if it applied in this setting, would imply that the number of lines spanned by P should be $\Omega(q^4)$, and the Szemerédi-Trotter theorem would imply that the number of q-rich lines cannot be more than $O(q)$.

In fact, the number of q-rich lines spanned by $\mathbb{F}_q^2$ matches an easy combinatorial upper bound that can be proved by using only the facts that each pair of points spans a unique line, and each pair of lines intersect in at most one point. Hence, it is not possible to prove incidence bounds for points and lines in $\mathbb{F}_q^2$ that beat the general combinatorial bounds, unless your proof is sensitive to the size of $|P|$ relative to the order or characteristic of $\mathbb{F}_q^2$.

The first non-trivial Szemerédi-Trotter and Beck's analogs over finite fields were proved by Bourgain, Katz, and Tao [11]. Suppose that p is an odd prime congruent to $3 \mod 4$, P is a set of n points in $\mathbb{F}_p^2$, and L is a set of n lines in $\mathbb{F}_p^2$, with $n < p^{2-\alpha}$ for some $\alpha > 0$. One result of [11] is that the number of incidences between P and L is bounded above by $n^{3/2-\epsilon}$, for some $\epsilon > 0$ depending on α. In this setting, the trivial combinatorial bound is $O(n^{3/2})$, and a true Szemerédi-Trotter analog would give $O(n^{4/3})$. There have been improvements to the bound of [11], most recently and dramatically by de Zeeuw and Stevens [57], but the best bounds in this setting remain far from $O(n^{4/3})$.

For large point sets ($|P| \subseteq \mathbb{F}_q^2, |P| > q$), more is known. Vinh proved that, for P a set of n points and L a set of n lines in $\mathbb{F}_q^2$, for an arbitrary prime power q, the number of incidences between P and L is bounded above by $n^2q^{-1} + nq^{1/2}$. When $n = q$, this matches the trivial bound of $O(n^{3/2})$, which is the best possible in this range of parameters. To see this, supppose that $\mathbb{F}_q$ has a subfield of size $q^{1/2}$, and take P and L to be the sets of all points and all lines defined over this subfield. When $n = q^{3/2}$, Vinh's bound is $O(q^{4/3})$. This also cannot be improved, which can be seen by taking P as an arbitrary Cartesian product of the appropriate size.

Chapter 4 contains a far-reaching generalization of Vinh's result. Instead of considering points and lines in $\mathbb{F}_q^2$, we show that a more general bound holds for balanced block designs. A special case of this bound is a generalization of Vinh's bound to incidences between points and k-flats in $\mathbb{F}_q^d$, for any $1 \leq k \leq d$.

We also give a stronger Beck's-type theorem than had been known previously, also in the setting of block designs. A special case of this bound is that a set of $(1+\varepsilon)q$ points in $\mathbb{F}_q^2$ spans $O(q^2)$ lines, for any fixed $\varepsilon > 0$ (the constant in the asymptotic notation depends on ε). We apply this Beck-type bound to improve a result of Iosevich, Rudnev, and Zhai on distinct triangle areas in the finite plane [39].

Chapter 2

Essential dimension and the number of flats spanned by a set of points

2.1 Introduction

Let P be a set of n points in a real or complex, finite-dimensional, affine space. We say that P spans a k-flat[1] Γ if Γ contains $k+1$ affinely independent points of P. Denote the number of k-flats spanned by P by f_k; in particular, $f_{-1} = 1$ and $f_0 = n$.

Our question is:

When does P span more k-flats than $(k-1)$-flats?

For $k = 1$, a complete answer to this question is given by a classic theorem of de Bruijn and Erdős [18]. This theorem is that either the number of lines spanned by P is at least n, or P is contained in a line; furthermore, equality is achieved only if $n-1$ points of P are collinear.

It might be tempting to conjecture that $f_k \geq f_{k-1}$ unless P is contained in a k-flat. This is easily seen to be false for $k = 2$, by considering a set of n points in $\mathbb{R}^3$, of which $n/2$ are incident to each of a pair of skew lines; in this case, $f_2 = n$ and $f_1 = (n/2)^2 + 2$. In 1986, Purdy [52] showed that either $n-1$ points of P lie on a plane or the union of a pair of skew lines, or $f_2 = \Omega(f_1)$.

To answer our question in higher dimensions, we introduce a new measure of the degeneracy of a point set with respect to affine subspaces. We say the *essential dimension* of a point set P is the minimum t such that there exists a set $\mathcal{G}$ of flats such that

1. P is contained in the union of the flats of $\mathcal{G}$,
2. each flat $\Gamma \in \mathcal{G}$ has dimension $\dim(\Gamma) \geq 1$, and
3. $\sum_{\Gamma \in \mathcal{G}} \dim(\Gamma) = t$.

For example, a point set that lies in the union of two skew lines has essential dimension 2. For any set of points P, we denote the essential dimension of P by $K(P)$, and we omit the argument if it is obvious from the context.

We additionally denote by g_i the maximum cardinality of a subset $P' \subseteq P$ such that the essential dimension of P' is at most i; i.e., $K(P') \leq i$.

We prove

[1] We refer to affine or projective subspaces as "flats".

Theorem 6. *For each k, there is a constant c_k such that the following holds. Let P be a set of n points in a finite dimensional real or complex affine geometry.*

1. *If $n = g_k$ (i.e., $K(P) \leq k$), then either $f_{k-1} > f_k$, or $f_{k-1} = f_k = 0$.*
2. *If $n - g_k > c_k$, then $f_k > f_{k-1}$.*

This theorem is a modification of a conjecture of Purdy [28]. A counterexample to Purdy's original conjecture for $k \geq 3$ was given by the author, Purdy, and Smith [47]; however, this counterexample left open the possibility that some variation on the conjecture (such as Theorem 6) could be true.

The case $k = 2$ of Purdy's conjecture was: if P is a set of sufficiently many points, then either P can be covered by two lines, or by a plane and a point, or P spans at least as many planes as lines (i.e., $f_2 \geq f_1$). This case of the conjecture appears in well-known collections of open problems in combinatorial geometry [12, 16], and has remained open until now. We give counterexamples to this conjecture in section 2.7, even showing that there are arbitrarily large point sets that cannot be covered by a plane and a point or by two lines such that $f_2 < (5/6)f_1 + O(1)$.

Also in Section 2.7, we investigate lower bounds on the values that may be taken by c_k in Theorem 6. In particular, we show that, even if we restrict our attention to arbitrarily large point sets, Theorem 6 does not hold for values of c_2 less than 4 or c_3 less than 11, and for larger k we show that c_k grows at least linearly with k. We further give a construction that we conjecture would show that c_k must grow at least exponentially with k, if we could properly analyze the construction in high dimensions.

Unlike the theorem of de Bruijn and Erdős mentioned above, Theorem 6 depends crucially on the underlying field. For example, consider the set P of all points in $\mathbb{F}_q^d$, where $\mathbb{F}_q$ is the finite field with q elements. The number of $(d-1)$-flats spanned by P is $\Theta(q^d)$, while the number of $(d-2)$-flats is $\Theta(q^{2(d-1)})$; however, no set of essential dimension $d-1$ contains more than q^{d-1} points.

Other than the result of Purdy for the case $k = 2$ mentioned above, the most relevant prior work on this question is a result of Beck [7], who proved that there is a constant

c'_k depending on k such that, either $f_k = \Omega(n^{k+1})$,[2] or a single hyperplane contains $c'_k n$ points. Hence, if P is a set of sufficiently many points, and no hyperplane contains more than a small, constant fraction of the points of P, then $f_k \geq f_{k-1}$. Considering the example of a set P of n points, n/k of which lie on each of k skew lines spanning $\mathbb{R}^{2k-1}$, shows that c'_k must be a decreasing function of k in this theorem.

The second claim (for $n - g_k \geq c_k$) of Theorem 6 is a consequence of the following asymptotic expression for the number of k-flats spanned by P.

Theorem 7. *Let P be a set of n points in a finite dimensional real or complex affine geometry.*

For $k < K = K(P)$,

$$f_k = \Theta\left(\prod_{i=0}^{k}(n - g_i)\right), \tag{2.1}$$

provided that $n - g_k \geq c_k$, for a constant c_k depending only on k.

For $k \geq K$,

$$f_k = O\left(\prod_{i=0}^{2(K-1)-k}(n - g_i)\right). \tag{2.2}$$

Claim 2 of Theorem 6 is an immediate consequence of expression (2.1) in Theorem 7. Theorem 7 is also a substantial generalization of a conjecture made by the author, Purdy, and Smith [47].

Recently, Do [19] independently found a different proof a special case of (2.1). In particular, Do shows that if $n - g_k = \Omega(n)$ then $f_k = \Omega(n^{k+1})$, for suitable choices of the implied constants.

Theorem 7 additionally implies an asymptotic version of a special case of a long-standing conjecture in matroid theory. Rota [30] conjectured that the sequence of the number of flats of each rank in any geometric lattice is unimodal, and Mason [48] proposed the stronger conjecture that the sequence is log-concave. We have

Corollary 8. *For $k < K$ such that $n - g_k \geq c_k$,*

$$f_k^2 = \Omega(f_{k-1}f_{k+1}).$$

[2] Here, and throughout the chapter, the constants hidden by asymptotic notation depend on k.

This follows immediately from Theorem 7 and the easy observation that $n - g_i \leq n - g_{i-1}$ for any i. Note that Corollary 8 applies only to real or complex affine geometries, and is also weaker than Rota's conjecture due the additional assumptions on P and the implied constant in the asymptotic notation.

We remark that the assumption that the underlying field is either the real or complex numbers is only used for the lower bound of Theorem 7; the proofs of claim 1 of Theorem 6 and the upper bound of Theorem 7 are independent of this assumption. Claim 2 of Theorem 6 and Corollary 8 both rely on the lower bound of Theorem 7, and hence are proved only for real and complex geometry.

2.1.1 A generalization of the Elekes-Tóth incidence bound

We also apply the idea of the essential dimension to generalize an incidence bound of Elekes and Tóth.

Incidence bounds are one of the most important tools in combinatorial geometry. The most famous such bound is the Szmerédi-Trotter theorem, stated below. We say that a k-flat is r-rich if it contains r or more points of P.

Theorem 9 (Szemerédi-Trotter)**.** *The number of r-rich lines spanned by a set of n points in the plane is bounded above by $O(n^2/r^3 + n/r)$.*

In order to generalize the Szemerédi-Trotter theorem to bound the number of r-rich k-flats for $k > 1$, we need to use some non-degeneracy assumption. Indeed, consider a set of r collinear points in $\mathbb{R}^3$; the number of r-rich planes is unbounded.

Elekes and Tóth [23] introduced one such degeneracy condition, and used it to obtain a strong bound on the number of incidences between points and planes. We say that a k-flat Λ is α-degenerate if at most $\alpha|P \cap \Lambda|$ points of P are contained in any $(k-1)$-flat contained in Λ.

Theorem 10 (Elekes, Tóth)**.** *For any $\alpha < 1$, the number of r-rich, α-degenerate 2-flats is bounded above by $O(n^3 r^{-4} + n^2 r^{-2})$.*

Elekes and Tóth also obtained a higher dimensional generalization of this bound, although in a weaker form.

Theorem 11 (Elekes, Tóth)**.** *For each $k > 2$, there is a constant β_k such that, for any $\alpha \leq \beta_k$, the number of r-rich, α-degenerate k-flats is bounded above by $O(n^{k+1}r^{-k-2} + n^k r^{-k})$.*

Recently, Do [19] used the idea of essential dimension to obtain an alternate generalization of Theorem 10. We say that a k-flat Λ is *essentially-α-degenerate* if largest subset of $P' \subseteq P \cap \Lambda$ with essential dimension $K(P') < k$ has at most $\alpha|P \cap \Lambda|$ points. Note that a set of points can be α-degenerate without being essentially-α-degenerate, but not visa-versa.

Theorem 12 (Do)**.** *For any k and any $\alpha < 1$, the number of essentially-α-degenerate, r-rich k-flats is bounded above by $O(n^{k+1}r^{-k-2} + n^k r^{-k})$.*

Notice that Theorems 11 and 12 are incomparable; Theorem 12 removes the constant β_k, but only bounds essentially-α-degenerate k-flats, rather than all α-degenerate k-flats.

In Section 2.8, we obtain the following proper generalization of Theorem 10.

Theorem 13. *For any k and any $\alpha < 1$, the number of α-degenerate, r-rich k-flats is bounded above by $O(n^{k+1}r^{-k-2} + n^k r^{-k})$.*

2.1.2 Organization of the chapter

Section 2.2 reviews basic facts of projective geometry and defines notation. Section 2.3 gives the proof of claim 1 of Theorem 6. Section 2.4 gives the proof of the upper bound of Theorem 7. Section 2.5 reviews some well-known consequences of the Szemerédi-Trotter theorem. Section 2.6 gives the proof of the lower bound of Theorem 7. Section 2.7 describes several new infinite families of point sets, that disprove Purdy's conjecture for $\mathbb{R}^3$, and establish lower bounds on the values that could be assumed by the constant c_k in Theorem 6. Section 2.8 contains the application of the idea of essential dimension to the incidence bound of Elekes and Tóth.

2.1.3 Acknowledgements

I thank George Purdy for suggesting the problem, and Abdul Basit, Zoltan Kiraly, Joe Malkevich, George Purdy, and Justin Smith for various helpful conversations and suggestions.

2.2 Preliminaries

In section 2.2.1, we review some basic facts of projective geometry, and fix the relevant notation. In section 2.2.2, we define some basic constructions, and list consisely the notation used for these constructions.

2.2.1 Projection

It suffices to prove Theorems 6 and 7 for sets of points in a finite dimensional projective geometry. Indeed, given a set of points in an affine geometry, we can add an empty hyperplane at infinity to obtain points in a projective geometry that determine the same lattice of flats.

In this section, we fix notation and review basic facts about projective geometry that we rely on in the proofs.

We denote by $\mathbb{P}^d$ the d-dimensional projective geometry over either $\mathbb{R}$ or $\mathbb{C}$. We refer to projective subspaces of $\mathbb{P}^d$ as flats.

The span of a set $X \subset \mathbb{P}^d$ is the smallest flat that contains X, and is denoted $\overline{X}$. Let Λ, Γ be flats of $\mathbb{P}^d$. We denote by $\overline{\Lambda, \Gamma}$ the span of $\Lambda \cup \Gamma$. It is a basic fact of projective geometry that

$$\dim(\overline{\Lambda, \Gamma}) + \dim(\Lambda \cap \Gamma) = \dim(\Lambda) + \dim(\Gamma). \tag{2.3}$$

Recall that $\dim(\emptyset) = -1$.

For a k-flat Λ, we define the projection from Λ to be the map

$$\pi_\Lambda : \mathbb{P}^d \setminus \Lambda \to \mathbb{P}^{d-k-1}$$

that sends a point p to the intersection of the $(k+1)$-flat $\overline{p,\Lambda}$ with an arbitrary $(d-k-1)$-flat disjoint from Λ. With a slight abuse of notation, for any set $X \subseteq \mathbb{P}^d$, we define $\pi_\Lambda(X)$ to be the image of $X \setminus \Lambda$ under projection from Λ.

For example, let Λ, Γ be flats in $\mathbb{P}^d$ of dimensions k and k', respectively. Then, $\pi_\Lambda(\Gamma)$ is defined to be the intersection of $\overline{\Gamma,\Lambda}$ with a $(d-k-1)$-flat Σ such that $\Lambda \cap \Sigma = \emptyset$. Together with equation (2.3), this implies that $\dim(\overline{\Sigma,\Lambda}) = d$, and, since we are in $\mathbb{P}^d$, we have also that $\dim(\overline{\Sigma,\Lambda,\Gamma}) = d$. Applying (2.3) two more times, we have

$$\begin{aligned}
\dim(\Sigma \cap \overline{\Lambda,\Gamma}) &= \dim(\Sigma) + \dim(\overline{\Lambda,\Gamma}) - \dim(\overline{\Lambda,\Gamma,\Sigma}), \\
&= \dim(\Sigma) + \dim(\Lambda) + \dim(\Gamma) - \dim(\Lambda \cap \Gamma) - \dim(\overline{\Lambda,\Gamma,\Sigma}), \\
&= k' - 1 - \dim(\Lambda \cap \Gamma).
\end{aligned}$$

In other words, the projection of a k'-flat through a k-flat in $\mathbb{P}^d$ is a $(k'-1-\dim(\Lambda\cap\Gamma))$-flat in $\mathbb{P}^{d-k-1}$, and so

$$\dim(\pi_\Lambda(\Gamma)) = \dim(\Gamma) - 1 - \dim(\Gamma \cap \Lambda). \tag{2.4}$$

2.2.2 Context and notation

For the remainder of the chapter, we fix a point set P of size $|P| = n$ in a finite dimensional real or complex projective space.

Recall that the *essential dimension* $K(Q)$ of a set Q of points is the minimum t such that there exists a set of flats, each of dimension 1 or more, the union of which contains Q, and whose dimensions sum to t. The proofs in sections 2.4 and 2.6 proceed primarily by isolating maximum size subsets of P having specified essential dimension.

We define $g_k(Q)$ to be the maximum size of a subset $Q' \subseteq Q$ such that $K(Q') \leq k$. We define $\mathcal{G}_k(Q)$ as a set of flats that satisfies the following conditions:

1. each flat in $\mathcal{G}_k$ has dimension at least 1,
2. $\sum_{\Gamma \in \mathcal{G}_k} \dim(\Gamma) \leq k$,
3. $|\cup_{\Gamma \in \mathcal{G}_k} \Gamma \cap Q| = g_k$,
4. $|\mathcal{G}_k| \leq |\mathcal{G}'_k|$ for any set $\mathcal{G}'_k$ that satisfies conditions 1,2, and 3.

In other words, $\mathcal{G}_k(Q)$ is a set of flats of minimum cardinality that contains a maximum cardinality set $Q' \subset Q$ with essential dimension $K(Q') \leq k$.

We further define the following functions on any point set Q:

$f_k(Q)$	the number of k-flats spanned by Q,
$\mathcal{F}_k(Q)$	the set of k-flats spanned by Q,
$f_k^{\sigma c}(Q)$	for $\sigma \in \{\leq, =, \geq\}$; the number of k-flats spanned by Q that each contain at most / exactly / at least c points of Q,
$\mathcal{F}_k^{\sigma c}(Q)$	the set of flats counted by $f_k^{\sigma c}(Q)$,
$\mathcal{G}(Q)$	$\mathcal{G}_{K(Q)}(Q)$.

The argument to any one of these functions will be omitted when it is clear from the context, in which case the argument will most often be P. This also applies to the projection operations described in section 2.2.1; for example, π_Γ is shorthand for $\pi_\Gamma(P)$, and denotes the projection of P from Γ.

Given a point set Q and a set of flats $\mathcal{F}$, we define the number of incidences between Q and $\mathcal{F}$ as

$$I(Q, \mathcal{F}) = |\{(p, \Gamma) \in Q \times \mathcal{F} \mid p \in \Gamma\}|.$$

2.3 Claim 1 of Theorem 6

In this section, we establish claim 1 of Theorem 6.

The results in this section are for weighted points. In particular, we assume the existence of a function $W : P \to \mathbb{R}$ such that $W(p) \geq 1$ for all $p \in P$.

Given such a weight function on the points of P, we extend it to flats and define related weight functions for projections of P as follows. The weight of a flat Λ is

$$W(\Lambda) = \sum_{p \in P \cap \Lambda} W(p).$$

The weight of a point $q \in \pi_\Gamma$ is

$$W_\Gamma(q) = \sum_{\substack{p \in P \mid \\ \pi_\Gamma(p) = q}} W(p).$$

Note that, for any flat Γ, we have

$$\sum_{q \in \pi_\Gamma} W_\Gamma(q) + W(\Gamma) = \sum_{p \in P} W(p).$$

The following simple lemma shows how to rewrite the sum of a function of the weights of the flats spanned by P in terms of the flats projected from each point $p \in P$.

Lemma 14. *For any function F and $k \geq 1$,*

$$\sum_{\Lambda \in \mathcal{F}_k} F(W(\Lambda)) = \sum_{p \in P} \sum_{\Lambda \in \mathcal{F}_{k-1}(\pi_p)} \frac{W(p) \cdot F(W_p(\Lambda) + W(p))}{W_p(\Lambda) + W(p)}.$$

Proof.

$$\begin{aligned}
\sum_{\Lambda \in \mathcal{F}_k} F(W(\Lambda)) &= \sum_{\Lambda \in \mathcal{F}_k} F(W(\Lambda)) \sum_{p \in P \cap \Lambda} \frac{W(p)}{W(\Lambda)}, \\
&= \sum_{p \in P} \sum_{\Lambda | p \in \Lambda} \frac{W(p) F(W(\Lambda))}{W(\Lambda)}, \\
&= \sum_{p \in P} \sum_{\Lambda \in \mathcal{F}_{k-1}(\pi_p)} \frac{W(p) \cdot F(W(\Lambda) + W(p))}{W(\Lambda) + W(p)}.
\end{aligned}$$

The last line uses the observation that the k-flats spanned by P and incident to p are in bijection with the $(k-1)$-flats spanned by π_p. □

The following lemma is the main claim of the section, from which claim 1 of Theorem 6 follows easily. We write $\mathbb{R}^+$ for the set of strictly positive real numbers.

Lemma 15. *Let $F : \mathbb{R}^+ \to \mathbb{R}^+$ be a non-increasing function. Let $k \geq K$, with $f_k \geq 1$. Then,*

$$\sum_{\Lambda \in \mathcal{F}_k} F(W(\Lambda)) < \sum_{\Lambda \in \mathcal{F}_{k-1}} F(W(\Lambda)).$$

Note that the conclusion $f_k < f_{k-1}$ follows by taking F to be the function that takes constant value 1.

Proof. We proceed by induction on K. In the base case, P is a collinear set of at least

2 points. Hence, for an arbitrary $p \in P$, we have

$$\sum_{\Lambda \in \mathcal{F}_1} F(W(\Lambda)) = F(W(P)) \leq F(W(p)) < \sum_{q \in P} F(W(q)),$$

which establishes the claim.

Now, assume that the lemma holds for $K' < K$ and arbitrary k. By Lemma 14, we have

$$\sum_{\Lambda \in \mathcal{F}_j} F(W(\Lambda)) = \sum_{p \in P} \sum_{\Lambda \in \mathcal{F}_{j-1}(\pi_p)} \frac{W(p) \cdot F(W_p(\Lambda) + W(p))}{W_p(\Lambda) + W(p)}, \tag{2.5}$$

for each of $j = k$ and $j = k - 1$.

Clearly, $K(\pi_p) \leq K - 1$. Indeed, let $p \in \Gamma \in \mathcal{G}$. Then π_p is contained in the union of $\pi_p(\Gamma)$ and $\pi_p(\Gamma')$ for $\Gamma' \in \mathcal{G} \setminus \Gamma$. Since $\dim(\pi_p(\Gamma)) = \dim(\Gamma) - 1$, this provides a witness that $K(\pi_p) \leq K - 1$.

Fix $p \in P$, and let

$$F_p(w) = \frac{W(p)F(w + W(p))}{w + W(p)},$$

defined for positive w. Since F is positive valued and nonincreasing, and $W(p) \geq 1$, we have that F_p is positive valued and nonincreasing. Hence, the induction hypothesis implies that

$$\sum_{\Lambda \in \mathcal{F}_{k-1}(\pi_p)} F_p(W_p(\Lambda)) < \sum_{\Lambda \in \mathcal{F}_{k-2}(\pi_p)} F_p(W_p(\Lambda)). \tag{2.6}$$

Together, (2.5) and (2.6) imply the conclusion of the lemma. □

2.4 Upper bound of Theorem 7

The main result of this section is Theorem 18, which is the upper bound of Theorem 7. Before proving the main result, we establish two lemmas on the set of k-flats spanned by P, for $k \geq K$.

Lemma 16. *Let $k \geq K$, and let $\Gamma \in \mathcal{F}_k$. Then, there is a set $\mathcal{A} \subseteq \mathcal{G}$ of $|\mathcal{A}| = k + 1 - K$ flats such that $\Lambda \subseteq \Gamma$ for each $\Lambda \in \mathcal{A}$.*

Proof. We first show that

$$\dim(\overline{\mathcal{A}}) \leq -1 + \sum_{\Lambda \in \mathcal{A}} (\dim(\Lambda) + 1). \tag{2.7}$$

We proceed by induction on $|\mathcal{A}|$. In the base case, $|\mathcal{A}| = 1$ and the claim holds. Suppose that $|\mathcal{A}| > 1$, and choose $\Lambda \in \mathcal{A}$ arbitrarily. By equation (2.3),

$$\begin{aligned}\dim(\overline{\mathcal{A}}) &= \dim(\Lambda) + \dim(\overline{\mathcal{A} \setminus \Lambda}) - \dim(\Lambda \cap \overline{\mathcal{A} \setminus \Lambda}),\\ &\leq \dim(\Lambda) + \dim(\overline{\mathcal{A} \setminus \Lambda}) + 1.\end{aligned}$$

The claim follows by the inductive hypothesis.

Let $\mathcal{A} \subseteq \mathcal{G}$ be the set of flats in $\mathcal{G}$ that are contained by Γ. We will show that $|\mathcal{A}| \geq k + 1 - K$.

Denote

$$\mathcal{G}_\Gamma = \{\Lambda \cap \Gamma \mid \Lambda \in \mathcal{G}\}.$$

Since each point of P is contained in some flat of $\mathcal{G}$, we have $\Gamma = \overline{\mathcal{G}_\Gamma}$. By (2.7),

$$k = \dim(\overline{\mathcal{G}_\Gamma}) \leq -1 + \sum_{\Delta \in \mathcal{G}_\Gamma} (\dim(\Delta) + 1). \tag{2.8}$$

If Δ is a flat contained in a flat Λ, then $\dim(\Delta) + 1 - \dim(\Lambda) \leq 1$, and if Δ is properly contained in Λ, then $\dim(\Delta) + 1 - \dim(\Lambda) \leq 0$. If $\Delta \in \mathcal{G}_\Gamma$ and $\Delta \in \mathcal{G}$, then $\Delta \in \mathcal{A}$. Hence,

$$\sum_{\Delta \in \mathcal{G}_\Gamma} (\dim(\Delta) + 1) - \sum_{\Lambda \in \mathcal{G}} \dim(\Lambda) \leq |\mathcal{A}|. \tag{2.9}$$

Since $\sum_{\Lambda \in \mathcal{G}} \dim(\Lambda) = K$ by definition, the conclusion of the lemma follows from inequalities (2.8) and (2.9). $\square$

Lemma 17. *Suppose $k \geq K$. Let $\mathcal{A} \subseteq \mathcal{G}$ such that $|\mathcal{A}| = k + 1 - K$ and $f_{k - \dim \overline{\mathcal{A}} - 1}(\pi_{\overline{\mathcal{A}}})$ is maximized. Let $k' = k - \dim \overline{\mathcal{A}} - 1$. Then,*

$$f_k = \Theta(f_{k'}(\pi_{\overline{\mathcal{A}}})).$$

Proof. Note that there is a natural bijection between flats of $\mathcal{F}_{k'}(\pi_{\overline{\mathcal{A}}})$ and flats of $\mathcal{F}_k$ that contain $\overline{\mathcal{A}}$. In particular, if $\Gamma \in \mathcal{F}_k$, then, by (2.4), we have

$$\dim(\pi_{\overline{\mathcal{A}}}(\Gamma)) = k - 1 - \dim(\Gamma \cap \overline{\mathcal{A}}) = k'.$$

In addition, $\overline{\pi_{\overline{\mathcal{A}}}(\Gamma), \mathcal{A}} = \Gamma$, so the map that sends each flat in $\mathcal{F}_k$ to its projection from $\mathcal{A}$ is invertible. Since the f_k is at least the number of flats in $\mathcal{F}_k$ that contain $\overline{\mathcal{A}}$, we have

$$f_k \geq f_{k'}(\pi_{\overline{\mathcal{A}}}).$$

On the other hand, by Lemma 16, for each k-flat $\Gamma \in \mathcal{F}_k$, there is at least one set $\mathcal{B} \subset \mathcal{G}$ with $|\mathcal{B}| = k + 1 - K$ such that $\Lambda \subset \Gamma$ for each $\Lambda \in \mathcal{B}$. Hence, we can define an injective function that maps each $\Gamma \in \mathcal{F}_k$ to an arbitrary pair $(\mathcal{B}, \Lambda)$ where $\mathcal{B}$ is a set as guaranteed by Lemma 16 and $\Lambda \in \mathcal{F}_{k-\dim \overline{A}-1}(\pi_{\overline{\mathcal{B}}})$ so that $\Gamma = \overline{\Lambda, \mathcal{B}}$. Since there are at most $\binom{K}{k+1-K} < 2^K \leq 2^k$ choices for $\mathcal{B}$, and $f_{k'}(\pi_{\overline{\mathcal{A}}}) \geq f_{k-\dim \overline{B}-1}(\pi_{\overline{\mathcal{B}}})$ by assumption, this shows that

$$f_k \leq 2^k f_{k'}(\pi_{\overline{\mathcal{A}}}),$$

which completes the proof of the lemma. $\square$

Next is the the main result of the section.

Theorem 18. *For* $0 \leq k \leq K - 1$*,*

$$f_k = O\left(\prod_{i=0}^{k}(n - g_i)\right). \tag{2.10}$$

For $k \geq K$*,*

$$f_k = O\left(\prod_{i=0}^{2(K-1)-k}(n - g_i)\right). \tag{2.11}$$

Proof. The proof is structured as follows. There is an outer induction on K. For a fixed K, we first prove inequality (2.11), and then use an induction on k to prove inequality (2.10).

The base case $k = 0$ and $K \geq 1$ is immediate, since $f_0 = n = n - g_0$ by definition.

Assume that inequalities (2.10) and (2.11) hold for all k when $K' < K$.

Suppose that $k \geq K$. By Lemma 16, either $|\mathcal{G}| \geq k+1-K$, or $f_k = 0$. If $f_k = 0$, then we're done, so suppose that $|\mathcal{G}| \geq k+1-K$.

By Lemma 17, there is a set $\mathcal{A} \subseteq \mathcal{G}$ with $|\mathcal{A}| = k+1-K$ such that $f_k = \Theta(f_{k'}(\pi_{\overline{\mathcal{A}}}))$, for $k' = k - \dim \overline{\mathcal{A}} - 1$.

Before bounding $f_{k'}(\pi_{\overline{\mathcal{A}}})$, we first make some simple observations about $\pi_{\overline{\mathcal{A}}}$. By definition, each point of $\pi_{\overline{\mathcal{A}}}$ is the image of one or more points that lie on flats of $\mathcal{G} \setminus \mathcal{A}$. Since $\dim(\Lambda) \geq \dim(\pi_{\overline{\mathcal{A}}}(\Lambda))$ for any flat Λ, the fact that the preimage of $\pi_{\overline{\mathcal{A}}}$ is contained the flats of $\mathcal{G} \setminus \mathcal{A}$ implies that

$$K(\pi_{\overline{\mathcal{A}}}) \leq \sum_{\Lambda \in \mathcal{G} \setminus \mathcal{A}} \dim(\Lambda) = K - \sum_{\Lambda \in \mathcal{A}} \dim(\Lambda).$$

Since $\sum_{\Lambda \in \mathcal{A}} \dim(\Lambda) \geq |\mathcal{A}| = k+1-K$, we have

$$K(\pi_{\overline{\mathcal{A}}}) \leq 2K - 1 - k.$$

In particular, $K(\pi_{\overline{\mathcal{A}}}) < K$, so we will be able to use the inductive hypothesis to bound $f_{k'}$.

Observe that the right sides of (2.10) and (2.11) are both bounded above by $O(\Pi_{i=0}^{K-1}(n - g_i))$. Hence, by the inductive hypothesis, we have that

$$\begin{aligned} f_{k'}(\pi_{\overline{\mathcal{A}}}) &= O\left(\Pi_{i=0}^{K(\pi_{\overline{\mathcal{A}}})-1}(|\pi_{\overline{\mathcal{A}}}| - g_i(\pi_{\overline{\mathcal{A}}}))\right), \\ &= O\left(\Pi_{i=0}^{2K-2-k}(|\pi_{\overline{\mathcal{A}}}| - g_i(\pi_{\overline{\mathcal{A}}}))\right). \end{aligned} \tag{2.12}$$

Note that $|\pi_{\overline{\mathcal{A}}}| - g_i(\pi_{\overline{\mathcal{A}}}) \leq n - g_i$ for each i. Indeed, the preimage of $\pi_{\overline{\mathcal{A}}} \cap \mathcal{G}_i(\pi_{\overline{\mathcal{A}}})$ has essential dimension at least i, so the preimage of $\pi_{\overline{\mathcal{A}}} \setminus (\pi_{\overline{\mathcal{A}}} \cap \mathcal{G}_i(\pi_{\overline{\mathcal{A}}}))$ provides a witness that $n - g_i \geq |\pi_{\overline{\mathcal{A}}}| - g_i(\pi_{\overline{\mathcal{A}}})$.

Together with (2.12), this completes the proof of (2.11).

Suppose now that $k \leq K-1$, and assume that inequality (2.10) holds for K and $k' < k$.

We claim that if P_1, P_2 is a partition of P, then

$$f_k \leq \sum_{i=-1}^{k} f_i(P_1) f_{k-i-1}(P_2). \tag{2.13}$$

To show this, we map $\mathcal{F}_k$ into $\bigcup_i (\mathcal{F}_i(P_1) \times \mathcal{F}_{k-i-1}(P_2))$. Let $\Gamma \in \mathcal{F}_k$, let $\Gamma_1 = \overline{P_1 \cap \Gamma}$, and let $\Gamma_2 = \overline{P_2 \cap \Gamma}$. Using equation (2.3) and the fact that $\dim(\Gamma_1 \cap \Gamma_2) \geq -1$, we have

$$\dim(\Gamma_2) \geq k - \dim(\Gamma_1) - 1.$$

Let $\Gamma_2' \subseteq \Gamma_2$ be a $(k - \dim(\Gamma_1) - 1)$-flat disjoint from Γ_1. Note that $\overline{\Gamma_1, \Gamma_2'} = \Gamma$. Also note that, if $\Gamma_1 = \Gamma$, then $\Gamma_2' = \emptyset$. Map Γ to the pair (Γ_1, Γ_2'). Since Γ is the unique k-flat spanned by Γ_1 and Γ_2', the map is injective, and so inequality (2.13) is established.

Let $P_1 = \cup_{\Gamma \in \mathcal{G}_k} (P \cap \Gamma)$, and let $P_2 = P \setminus P_1$. By inequality (2.13),

$$\begin{aligned} f_k &\leq \sum_{i=-1}^{k} f_i(P_1) f_{k-i-1}(P_2), \\ &\leq (k+2) \max_{-1 \leq i \leq k} f_i(P_1) f_{k-i-1}(P_2). \end{aligned} \tag{2.14}$$

Since $|P_2| = n - g_k$, we have

$$f_{k-i-1}(P_2) \leq (n - g_k)^{k-i} \leq \prod_{j=0}^{k-i-1} (n - g_{k-j}). \tag{2.15}$$

For $i < k$, the inductive hypothesis implies

$$f_i(P_1) = O\left(\prod_{j=0}^{i} (|P_1| - g_j(P_1))\right) = O\left(\prod_{j=0}^{i} (n - g_j)\right). \tag{2.16}$$

For $i = k$, inequality (2.11) implies

$$f_k(P_1) = O\left(\prod_{j=0}^{k-2} (|P_1| - g_j(P_1))\right) = O\left(\prod_{j=0}^{k} (n - g_j)\right). \tag{2.17}$$

With an appropriate choice of the constants hidden in the asymptotic notation, this completes the proof of inequality (2.10). □

2.5 Known results in the plane

In order to prove the lower bounds of Theorem 7, we will use two known consequences of the Szemerédi-Trotter theorem.

The Szemerédi-Trotter theorem was proved for real geometry by Szemerédi and Trotter [59], and proved for complex geometry by Tóth [62], and, using a different method, by Zahl [67].

Theorem 19. *[Szemerédi-Trotter] For any t,*

$$f_1^{\geq t} = O(n^2/t^3 + n/t).$$

Theorem 20 was proved by Beck [7] when the underlying field is the real numbers, and the idea of Beck's proof is easily adapted to use Theorem 19.

Theorem 20 (Beck)**.** *There is a constant c_b such that*

$$f_1^{\leq c_b} = \Omega(n(n - g_1)).$$

Proof. Let $0 < c_1 < 1$ be a constant to fix later. Counting pairs of points of P that are on lines that contain between c_b and $c_1 n$ points of P, we have

$$\begin{aligned}
\sum_{t=c_b}^{c_1 n} f_1^{=t} t^2 &= \sum_{t=c_b}^{c_1 n} t^2 (f_1^{\geq t} - f_1^{\geq t+1}), \\
&= \sum_{t=c_b}^{c_1 n} t^2 f_1^{\geq t} - \sum_{t=c_b+1}^{c_1 n+1} (t-1)^2 f_1^{\geq t}, \\
&= O\left(\sum_{t=c_b}^{c_1 n} t f_1^{\geq t} \right).
\end{aligned}$$

Applying Theorem 19, for appropriate choices of c_b and c_1 we have

$$O\left(\sum_{t=c_b}^{\sqrt{n}} t f_1^{\geq t} \right) = O\left(\sum_{t=c_b}^{\sqrt{n}} n^2/t^2 \right) \leq n^2/10,$$

and

$$O\left(\sum_{t=\sqrt{n}}^{c_1 n} t f_1^{\geq t}\right) = O\left(\sum_{t=\sqrt{n}}^{c_1 n} n\right) \leq n^2/10.$$

Hence, either at least $n^2/4$ pairs of points are on lines that each contain at most c_b points, or at least $n^2/4$ pairs of points are on lines that contain at least $c_1 n$ points. In the first case, $f_1^{\leq c_b} \geq n^2/(4c_b^2)$, and the theorem is proved. Hence, we suppose that $g_1 > c_1 n$.

Let ℓ be a line incident to g_1 points of P, and let P' be a set of $\min(g_1, n - g_1)$ points that are not incident to ℓ. Let L be the set of lines that contain one point of $P \cap \ell$ and at least one point of P'. Since each point of P' is incident to g_1 lines of L, we have

$$\sum_{l \in L} |P' \cap l| = |P'| g_1.$$

Since each ordered pair of distinct points in P' is incident to at most one line of L, we have

$$\sum_{l \in L} (|P' \cap l|^2 - |P' \cap l|) \leq |P'|^2 - |P'|.$$

By Cauchy-Schwarz,

$$\sum_{l \in L} |P' \cap l|^2 \geq \frac{\left(\sum_{\ell \in L} |P' \cap \ell|\right)^2}{|L|} = \frac{|P'|^2 g_1^2}{|L|}.$$

Combining these and rearranging, we have

$$|L| \geq \min(|P'| g_1, g_1^2) = \Omega(n(n - g_1)).$$

It remains to show that a constant portion of the lines of L each contain at most c_b points of P. Let P'' be the set of $n - g_1$ points of P that are not incident to ℓ. Each pair of points of P'' is incident to at most 1 line of L, hence the expected number of pairs of points of P'' on a randomly chosen line of L is at most $\binom{n-g_1}{2}|L|^{-1} = O(1)$. Markov's inequality implies that at least half of the lines of L are each incident to at most twice the expected number of points of P'', and the conclusion of the theorem follows. ☐

Theorem 21 is a variant of the "weak Dirac" theorem, proved independently by Beck [7], and by Szemerédi and Trotter [59].

Theorem 21 (Weak Dirac)**.** *There is a constant c_d such that, if P does not include $c_d n$ collinear points, then there is a subset $B \subseteq P$ with $|B| = \Omega(|P|)$ such that each point in B is incident to at least $\Omega(n)$ lines spanned by P.*

Proof. By Theorem 20, if no line contains $c_d n$ points of P, then P spans $\Omega(n^2)$ lines. Since no point is incident to more than n such lines, there must be $\Omega(n)$ points each incident to $\Omega(n)$ of these lines. ☐

2.6 Lower bound of Theorem 7

In this section, we prove Theorem 24, which gives the lower bound of Theorem 7.

We will need the following consequence of the minimality of $\mathcal{G}_k$.

Lemma 22. *For arbitrary k, let $\mathcal{A} \subseteq \mathcal{G}_k$, with $|\mathcal{A}| \geq 2$, and let Λ be an arbitrary flat. Then*

$$\sum_{\Gamma \in \mathcal{A}} \dim(\Gamma \cap \Lambda) < \dim(\Lambda).$$

Proof. Label the flats in $\mathcal{A}$ as $\Gamma_1, \ldots, \Gamma_{|\mathcal{A}|}$. Let $\Lambda_i = \overline{\Gamma_1, \ldots, \Gamma_i, \Lambda}$, with $\Lambda_0 = \Lambda$.

We claim that

$$\dim(\Lambda_i) \leq \dim(\Lambda) - \sum_{j=1}^{i} \dim(\Gamma_j \cap \Lambda) + \sum_{j=1}^{i} \dim(\Gamma_j). \tag{2.18}$$

The proof of (2.18) is by induction on i. In the base case, $i = 0$ and the claim is trivial.

Suppose (2.18) holds for $i' < i$. Then, applying equation (2.3),

$$\dim(\overline{\Gamma_i, \Lambda_{i-1}}) + \dim(\Gamma_i \cap \Lambda_{i-1}) = \dim(\Gamma_i) + \dim(\Lambda_{i-1}), \text{ so}$$
$$\dim(\Lambda_i) + \dim(\Gamma_i \cap \Lambda) \leq \dim(\Gamma_i) + \dim(\Lambda_{i-1}).$$

Inequality (2.18) follows by the inductive hypothesis.

Hence,

$$\dim(\overline{\mathcal{A}}) \leq \dim(\Lambda_{|\mathcal{A}|}) \leq \dim(\Lambda) + \sum_{\Gamma \in \mathcal{A}} \dim(\Gamma) - \sum_{\Gamma \in \mathcal{A}} \dim(\Gamma \cap \Lambda). \tag{2.19}$$

If we suppose that $\dim(\Lambda) \leq \sum_{\Gamma \in \mathcal{A}} \dim(\Gamma \cap \Lambda)$, then (2.19) implies that $\dim(\overline{\mathcal{A}}) \leq \sum_{\Gamma \in \mathcal{A}} \dim(\Gamma)$. Hence, we can reduce the size of $\mathcal{G}_k$ by replacing $\mathcal{A}$ by $\overline{\mathcal{A}}$, which contradicts the minimality of $\mathcal{G}_k$. $\square$

We use Lemma 22 to control the projection of the points contained in flats of $\mathcal{G}_k$ from a point in P that is not contained in a flat of $\mathcal{G}_k$.

Lemma 23. *Let $k < K$, let $A = \cup_{\Gamma \in \mathcal{G}_k} \Gamma \cap P$, and let $p \in P \setminus A$. Then, for $0 \leq i \leq k-1$,*

$$g_i(\pi_p(A)) \leq g_i(A) + k^2, \tag{2.20}$$

$$|\pi_p(A)| \geq |A| - k^2. \tag{2.21}$$

Proof. We first prove (2.20). Let $\Lambda \in \mathcal{G}_i(\pi_p(A))$, and let Λ' be the preimage of Λ under π_p; note that $\dim(\Lambda') = \dim(\Lambda) + 1$.

Let

$$\mathcal{L}(\Lambda) = \{\Gamma \cap \Lambda' \mid \Gamma \in \mathcal{G}_k, \dim(\Gamma \cap \Lambda') \geq 1\}.$$

Note that, since $p \notin A$, no flat in $\mathcal{L}(\Lambda)$ can contain p. Hence, if $\mathcal{L}(\Lambda)$ contains a single flat Γ, then $\dim(\Gamma) < \dim(\Lambda') = \dim(\Lambda) + 1$. On the other hand, if $|\mathcal{L}(\Lambda)| \geq 2$, then Lemma 22 implies that $\sum_{\Gamma \in \mathcal{L}(\Lambda)} \dim(\Gamma) < \dim(\Lambda') = \dim(\Lambda) + 1$. In either case, the flats of $\mathcal{L}(\Lambda)$ contain at most $g_{\dim(\Lambda)}(A)$ points of A. Since Λ is the projection of the points on flats of $\mathcal{L}(\Lambda)$ together with at most one point on each flat in $\mathcal{G}_k$ that does not intersect Λ' in at least a line, we have that $|\Lambda \cap \pi_p(A)| \leq g_{\dim \Lambda}(A) + k$. Note that, since $K((\mathcal{G}_i \cup \mathcal{G}_j) \cap P) \leq i + j$, we have that $g_i + g_j \leq g_{i+j}$ for any i, j. In particular,

$$\sum_{\Lambda \in \mathcal{G}_i(\pi_p(A))} g_{\dim \Lambda}(A) \leq g_i(A).$$

Hence, we have

$$g_i(\pi_p(A)) = \sum_{\Lambda \in \mathcal{G}_i(\pi_p(A))} |\Lambda \cap \pi_p(A)| \leq g_i(A) + ik,$$

which completes the proof of (2.20).

It remains to prove (2.21). Let $\Gamma, \Gamma' \in \mathcal{G}_k$. Since $\Gamma \cap \Gamma' = \emptyset$, we have $\dim(\Gamma' \cap \overline{\Gamma, p}) \leq 0$. Hence, for each such pair of flats $\Gamma, \Gamma' \in \mathcal{G}_k$, there is at most one pair $q \in \Gamma, q' \in \Gamma'$ of points such that $\pi_p(q) = \pi_p(q')$. In addition, each line incident to p intersects each flat of $\mathcal{G}_k$ in at most one point, since otherwise p would be contained in that flat. Hence, the number of pairs of points $q, q' \in A$ such that $\pi_p(q) = \pi_p(q')$ is at most the number of pairs of flats in $\mathcal{G}_k$, which proves (2.21). □

We now proceed to the main result of the section. Theorem 24 is slightly stronger than the lower bound of Theorem 7, to facilitate its inductive proof.

Theorem 24. *For $0 \leq k < K$, there are constants c_l, c_k such that*

$$f_k^{\leq c_l} = \Omega\left(\prod_{i=0}^{k}(n - g_i)\right),$$

provided that $n - g_k \geq c_k$.

Proof. The proof is by induction on k. The case $k = 1$ is Theorem 20.

Let

$$A = \bigcup_{\Gamma \in \mathcal{G}_k} \Gamma \cap P,$$

$$B = P \setminus A.$$

Note that $|A| = g_k$ and $|B| = n - g_k \geq c_k$.

Let $c_1 < 1$ be a strictly positive constant to fix later. Let k' be the least integer such that $|A| - g_{k'}(A) < c_1|B| = c_1(n - g_k)$.

If $k' < k$, then no line contains $c_1|B|$ points of B. Indeed, if ℓ is such a line, then $\mathcal{G}_{k'} \cup \ell$ contains $g_{k'} + c_1|B| > g_k$ points of P, which is a contradiction, since the sum of the dimensions of the flats of $\mathcal{G}_{k'} \cup \ell$ is $k' + 1 \leq k$.

If $k' = k$, let $B' = B$. Otherwise, by Theorem 21 (assuming $c_1 < c_d$), there is a set $B' \subseteq B$ with $|B'| = \Omega(|B|)$ such that each point of B' is incident to $\Omega(|B|)$ lines spanned by B.

Fix $p \in B'$ arbitrarily.

We claim that, for $0 \leq i \leq k - 1$,

$$|\pi_p| - g_i(\pi_p) = \Omega(n - g_i). \tag{2.22}$$

Recall that $0 \leq i \leq k-1$, and $k' \leq k$, and hence, it will suffice to consider the cases that $i < k'$ and $k' \leq i \leq k - 1$.

First, suppose that $i < k'$. Since k' is the least integer such that $|A| - g_{k'}(A) < c_1(n - g_k)$, we have that $|A| - g_i(A) \geq c_1(n - g_k)$. Using this fact, together Lemma 23, we have

$$\begin{aligned} n - g_i &= n - g_k + g_k - g_i, \\ &\leq (c_1^{-1} + 1)(|A| - g_i(A)), \\ &\leq (c_1^{-1} + 1)(|\pi_p(A)| - g_i(\pi_p(A)) + 2k^2), \\ &= O(|\pi_p(A)| - g_i(\pi_p(A))). \end{aligned} \tag{2.23}$$

In the last line of the above derivation, we require $|\pi_p(A)| - g_i(\pi_p(A)) > 0$. This holds if $|A| - g_i(A) > 2k^2$, which holds if $c_1 c_k > 2k^2$. Hence, we require $c_1 c_k > 2k^2$.

Since $\pi_p(A)$ is a subset of π_p, we have

$$|\pi_p| - g_i(\pi_p) \geq |\pi_p(A)| - g_i(\pi_p(A)).$$

Combined with (2.23), this is inequality (2.22).

Now, suppose that $k' \leq i \leq k - 1$.

Let $\Gamma \in \mathcal{G}_{k'}(A)$. Note that $|\overline{p, \Gamma} \cap B| < c_1|B|$. If this were not the case, then $\overline{p, \Gamma} \cup \mathcal{G}_{k'} \setminus \Gamma$ would have total dimension $k' + 1 \leq k$, and would contain at least

$g_{k'}(A) + c_1|B| > |A| = g_k$ points. Since $\mathcal{G}_{k'}$ contains at most $k' \leq k-1$ distinct flats, and the remaining points of A contribute at most $|A| - g_{k'}(A) < c_1|B|$ points to $|\pi_p(A) \cap \pi_p(B)|$, we have that $|\pi_p(A) \cap \pi_p(B)| \leq kc_1|B|$. Hence,

$$|\pi_p| \geq |\pi_p(A)| + |\pi_p(B)| - kc_1|B|. \tag{2.24}$$

Note that $g_i(\pi_p) \leq g_{i+1} \leq g_k$. Hence, by inequality (2.21) of Lemma 23, we have that $|\pi_p(A)| - g_i(\pi_p) \geq |\pi_p(A)| - g_k \geq -O(1)$. Combining this with inequality (2.24) and the assumption that $|B| > c_k$, we have

$$\begin{aligned} |\pi_p| - g_i(\pi_p) &\geq |\pi_p(A)| + |\pi_p(B)| - kc_1|B| - g_i(\pi_p), \\ &\geq c_d|B| - O(1) - kc_1|B|, \\ &= \Omega(|B|), \end{aligned}$$

for appropriate choices of c_1, c_k. Since $i \geq k'$, we have that $|B| = \Omega(n - g_i)$, and hence, this finishes the proof of inequality (2.22).

The inductive hypothesis applied to π_p, along with (2.22), implies that

$$f_{k-1}^{\leq O(1)}(\pi_p) = \Omega\left(\prod_{i=0}^{k-1}(n - g_i)\right). \tag{2.25}$$

Hence, each point in B' is incident to $\Omega\left(\prod_{i=0}^{k-1}(n-g_i)\right)$ flats of dimension k that are spanned by P. Since the preimage of a point $q \in \pi_p$ may include many points of P, it remains to show that a substantial portion of these flats each contain at most $c_l(k)$ points of P.

Let c_2 be a large constant, to be fixed later. Let $C \subset \pi_p$ be the set of points in π_p such that each point in C is the image of at least c_2 points in P under projection from p. Since each line incident to p is incident to at most one point on each flat $\Gamma \in \mathcal{G}_k$, each point of $\pi_p(A)$ has multiplicity at most $k < c_2$. Hence, $|C| \leq c_2^{-1}|B| = c_2^{-1}(n - g_k)$.

Let $q \in C$. By Theorem 18,

$$f_{k-2}(\pi_{\overline{q,p}}) = O\left(\prod_{i=0}^{k-2}(n-g_i)\right), \tag{2.26}$$

and this is an upper bound on the number of incidences between q and $(k-1)$-flats spanned by π_p.

The total number of $(k-1)$-flats spanned by π_p that are incident to some point in C is bounded above by the number of incidences between points in C and flats in $\mathcal{F}_{k-1}(\pi_p)$. Summing expression (2.26) over the points of C, and using the fact that $n - g_k < n - g_{k-1}$, the number of these incidences is

$$I(C, \mathcal{F}_{k-1}(\pi_p)) = O\left(c_2^{-1}\prod_{i=0}^{k-1}(n-g_i)\right). \tag{2.27}$$

By setting c_2 to be sufficiently large, we can ensure that the right side of (2.27) is smaller than the right side of (2.25). Hence, we can subtract from the right side of (2.25) the number of $k-1$ flats spanned by π_p that contain a point of C to obtain

$$I(p, \mathcal{F}_k^{\leq c_2 c_l(k-1)}) = \Omega\left(\prod_{i=0}^{k-1}(n-g_i)\right). \tag{2.28}$$

This bound applies for each of the $\Omega(n-g_k)$ points in B', and hence (setting $c_l(k) = c_2 c_l(k-1)$)

$$I(B', \mathcal{F}_k^{\leq c_l}) = \Omega\left(\prod_{i=0}^{k}(n-g_i)\right). \tag{2.29}$$

Since each of the flats of $\mathcal{F}_k^{\leq c_l}$ accounts for at most c_l of these incidences, dividing the right side by c_l immediately gives the claimed lower bound on $f_k^{\leq c_l}$. $\square$

2.7 Constructions

In this section, we give several constructions that give lower bounds on the possible values that could be taken by c_k in Theorem 6. We are in fact interested primarily in infinite families of examples for each k. Hence, for this section, we define c_k to be a function of k as follows.

Definition 25. *The constant c_k is the minimum t such that the following holds for all sufficiently large n. If P is a set of n points in $\mathbb{R}^d$ or $\mathbb{C}^d$, then either*

1. $n - g_k \leq t$, *or*
2. $f_k > f_{k-1}$.

Note that this definition includes the hypothesis that n is sufficiently large, which is absent in Theorem 6. Because of this aditional hypothesis, in order to show lower bounds of the form $c_k \geq t$, we find infinite families of point sets S_n, such that for each S_n we have $|S_n| = n$, $f_k(S_n) \leq f_{k-1}(S_n)$, and $n - g_k(S_n) = t$.

To summarize the results on c_k in this section, we show that c_k increases monotonically (subsection 2.7.1), that $c_k \geq k - O(1)$ (subsection 2.7.2), and that $c_2 \geq 4$ and $c_3 \geq 11$ (subsection 2.7.3). Also in subsection 2.7.3, we give strong counterexamples to the conjecture of Purdy mentioned in the introduction.

In subsection 2.7.2, we present a construction that we conjecture would show that $c_k \geq 2^{k-1}$ if it were successfully analyzed, but are unable to fully analyze the construction in higher dimensions.

All of the constructions in this section are based on the same basic idea, presented in subsection 2.7.1.

2.7.1 Basic construction, and monotonicity

All of the constructions described in this section follow the same basic plan. We start with a finite set S of points having some known properties, then carefully select an origin point, and place a line L, containing a large number of points of P, perpendicular to the hyperplane containing S and incident to the selected origin point. This construction, along with its key properties, is described in Lemma 26.

Lemma 26. *Let S be a set of n points in $\mathbb{R}^d$, all contained in the hyperplane H_0 defined by $x_1 = 0$. Denote by $f_k^o(S)$ the number of k-flats spanned by S that are incident to the origin, and by $f_k^{\overline{o}}(S)$ the number of k-flats spanned by S that are not incident to the origin; we define $f_0^o = 0$. Let L be a set of $m \geq 2$ collinear points contained in the line*

ℓ_0 defined by the equations $x_i = 0$ for $i \neq 1$, and stipulate that the origin is not included in L. Let $P = S \cup L$. Then, for each $0 < k < d$,

$$f_k(P) = mf_{k-1}^{\bar{o}}(S) + f_{k-1}^{o}(S) + f_k(S) + f_k(L). \tag{2.30}$$

Proof. Let $\Gamma \in \mathcal{G}_k(P)$. If Γ contains the origin and another point in ℓ_0, then Γ contains ℓ_0 and hence contains each point of L. In this case, $\dim(\Gamma \cap H_0) = k-1$, and there are $f_{k-1}^{o}(S)$ such flats spanned by S. If Γ contains exactly one point of L, then Γ does not contain the origin, and $\dim(\Gamma \cap H_0) = k-1$. Since there are m choices for the point in L, the number of such flats is $mf_{k-1}^{\bar{o}}(S)$. We also have those k-flats that are spanned individually by S or L. $\square$

Given an example that shows that $c_k \geq t$ for some t, Lemma 26 can be used to create an equally strong example for c_{k+1}, which implies that the sequence $c_2, c_3, \ldots$ is monotonic.

Corollary 27. *The sequence $c_2, c_3, \ldots$ increases monotonically.*

Proof. Let $1 < k < d$ and $c \geq 1$, and let S be a set of points in $\mathbb{R}^d$, such that $f_k(S) < f_{k-1}(S)$, and such that $|S| - g_k(S) = c_k$. Embed S in the hyperplane defined by $x_1 = 0$ in $\mathbb{R}^{d+1}$, so that no flat spanned by S is incident to the origin. Let L be a set of m points contained in the line $x_i = 0$ for $i \neq 1$, and not including the origin. Then, by Lemma 26, we have

$$f_{k+1}(P) = mf_k(S) + f_{k+1}(S) < mf_{k-1}(S) + f_k(S) = f_k(P),$$

for m sufficiently large.

In addition, since $|L|$ is much larger than $|S|$, we may assume that $\mathcal{G}_{k+1}(P)$ contains L. Since the origin is generic relative to the flats spanned by S, the number of points of S in a $j+1$ flat that contains the origin is bounded by the number of points in a j flat. Hence, $\mathcal{G}_{k+1}(P)$ is the union of the line that contains L and $\mathcal{G}_k(S)$, and hence $|P| - g_{k+1}(P) = |S| - g_k(S) = c_k \leq c_{k+1}$. $\square$

2.7.2 Constructions for arbitrary dimensions

We describe two constructions that work for any sufficiently large k. The first uses a hypercube as the set S in the construction of Lemma 26, and the second uses a cross-polytope as S. We are unable to fully analyze the hypercube example in arbitrary dimensions, but conjecture that a complete analysis would show that $c_k \geq 2^{k-1}$. The cross-polytope example shows that $c_k \geq k - O(1)$.

Hypercube construction. We use Lemma 26 to describe an infinite family of sets of points, with an infinite number of members for each $k \geq 2$. In particular, S_n^k, for $n \geq 2^{k+1}$, is a set of n points in $\mathbb{R}^{k+1}$ such that $n - g_k(S_n^k) = 2^{k-1}$. We conjecture that $f_k(S_n^k) < f_{k-1}(S_n^k)$ for all k. Proving this conjecture would show that $c_k \geq 2^{k-1}$. Analyzing the construction for large k is related to (though possibly easier than) the open problem of characterizing the set of flats spanned by the vertices of the hypercube $[-1,+1]^d$ in $\mathbb{R}^d$ (see [1]). It is easy, though tedious, to analyze the construction in low dimensions; however, different, specific constructions for $k = 2, 3$ give better bounds on c_k for $k \leq 4$.

Let $S_n^k = C^k \cup L$, where $C^k = (0, \pm 1, \ldots, \pm 1)$ is the set of vertices of a k-dimensional hypercube, and L is the set of $m = n - 2^k$ collinear points with coordinates $(i, 0, \ldots, 0)$ for $i \in [1, n - 2^k]$.

We claim that $g_k(S_n^k) = m + 2^{k-1}$. That $g_k \geq m + 2^{k-1}$ follows by considering the union of L and a $(k-1)$-dimensional face of C^k. To show that $g_k \leq m + 2^{k-1}$, we show that $g_{k-1}(C^k) = 2^{k-1}$; the claim on $g_k(S_k)$ follows as an immediate consequence, since $\mathcal{G}_k$ must contain L.

We show by induction that the intersection of a j-flat with C^k contains at most 2^j points, for any $j \leq k$. Note that $C^k = C_{-1}^{k-1} \cup C_1^{k-1}$, where C_i^{k-1} (for $i \in \{-1, 1\}$) is the set of vertices of a $(k-1)$-dimensional hypercube in the $(d-2)$-flat H_i defined by $x_0 = i$. Let Γ be a flat of dimension $\dim(\Gamma) = j$. Either Γ is contained in H_{-1}, or is contained in H_1, or intersects each of H_{-1} and H_1 in a $(j-1)$-flat. Assuming the inductive hypothesis that the intersection of a j'-flat with C^{k-1} contains at most $2^{j'}$ points, it follows that Γ contains at most 2^j points of C^k. Since the sum of the dimensions of flats in $\mathcal{G}_{k-1}(C^k)$ is k, it follows that $g_{k-1}(C^k) \leq 2^{k-1}$.

For $k = 2$ and $k = 3$, an exhaustive enumeration of the flats spanned by C^k is easy to perform by hand, and, for $k = 3$, yields

$$f_1^o(C^3) = 4,$$
$$f_1^{\overline{o}}(C^3) = 24,$$
$$f_2^o(C^3) = 6,$$
$$f_2^{\overline{o}}(C^3) = 14.$$

Together with a similar count for $k = 2$, and an application of Lemma 26, we have

$$f_1(S_2) = 4m + 7,$$
$$f_2(S_2) = 4m + 3,$$
$$f_2(S_3) = 24m + 24,$$
$$f_3(S_3) = 14m + 7.$$

Hence, our conjecture holds for these cases.

Cross-polytope construction. We describe a family of sets T_n^j of points for $j \geq 2$ and n sufficiently large. The set T_n^j is a set of $n = m+6j$ points in $\mathbb{R}^{3j+1}$ such that, assuming m is sufficiently large, then $f_{2j+2} < f_{2j+1} < f_{2j}$. Furthermore, $n - g_{2j+2} = 2j - 2$ and $n - g_{2j+1} = 2j$. Taking $k = 2j + 2$ in this construction shows that $c_k \geq k - 4$ for even $k \geq 6$, and taking $k = 2j + 1$ shows that $c_k \geq k - 1$ for odd $k \geq 5$.

Let $D = D^{3j}$ be the vertices of a $3j$-dimensional cross-polytope in $\mathbb{R}^{3j+1}$, centered at the origin, contained in the hyperplane $x_1 = 0$. In particular, the $6j$ vertices of D are of the form $(0, \ldots, 0, \pm 1, 0, \ldots, 0)$, where the nonzero entries occur for some vertex in all but the first coordinate. We use D as the set S in the construction of Lemma 26, so $T_n^j = D \cup L$, where L is a set of m points in the line $x_i = 0$ for $i \neq 1$. We will assume that m is large relative to $6j$.

We first show that $f_{2j+2} < f_{2j+1} < f_{2j}$. Let $v \in D$. If a flat Γ contains v and

$-v$, then Γ contains the origin. Hence, the i-flats spanned by D that don't contain the origin each contain at most one of $v, -v$. Since the non-opposite vertices of D are linearly independent, an i-flat contains at most $i+1$ of them, and so $f_i^{\overline{o}}(D)$ is equal to the number of ways to choose $i+1$ non-opposite vertices from D, which is $2^{i+1}\binom{3j}{i+1}$. Hence, we have

$$f_i^{\overline{o}}(D) = 2^{i+1}\binom{3j}{i+1} = 2^i\binom{3j}{i} \cdot 2\frac{3j-i}{i+1} = f_{i-1}^{\overline{o}}(D) \cdot 2\frac{3j-i}{i+1}.$$

Hence, if $(3j-i)/(i+1) < 1/2$, then $f_i^{\overline{o}}(D) < f_{i-1}^{\overline{o}}(D)$. This holds if $i \geq 2j$. Applying Lemma 26, and using the assumption that m is sufficiently large, we have

$$f_{2j+2} = f_{2j+1}^{\overline{o}}(D)m + O(1) < f_{2j}^{\overline{o}}(D)m + O(1) = f_{2j+1},$$
$$f_{2j+1} = f_{2j}^{\overline{o}}(D)m + O(1) < f_{2j-1}^{\overline{o}}(D)m + O(1) = f_{2j}.$$

Now we show that $n - g_{2j+2} = 2j-2$ and $n - g_{2j+1} = 2j$. In particular, we show that $g_i(D) = 2i$; since m is large, $g_{i+1} = m + g_i(D)$, and so $n - g_{i+1} = 6j - g_i(D) = 6j - 2i$, from which the claims easily follow.

Let Γ be an i-flat, for $i \geq 1$. If Γ contains the origin, then it is a linear subspace and hence contains at most i linearly independent vectors, and hence at most $2i$ vertices of D. If Γ does not contain the origin, then it contains at most $i+1$ linearly independent vectors, and does not contain any pair $v, -v \in D$; in this case, Γ contains at most $i+1$ vertices. In either case, Γ contains at most $2i$ vertices. Since the sum of the dimensions of the flats in $\mathcal{G}_i(D)$ is i, it's clear from this that $g_i(D) = 2i$.

2.7.3 Stronger constructions for $k = 2, 3$

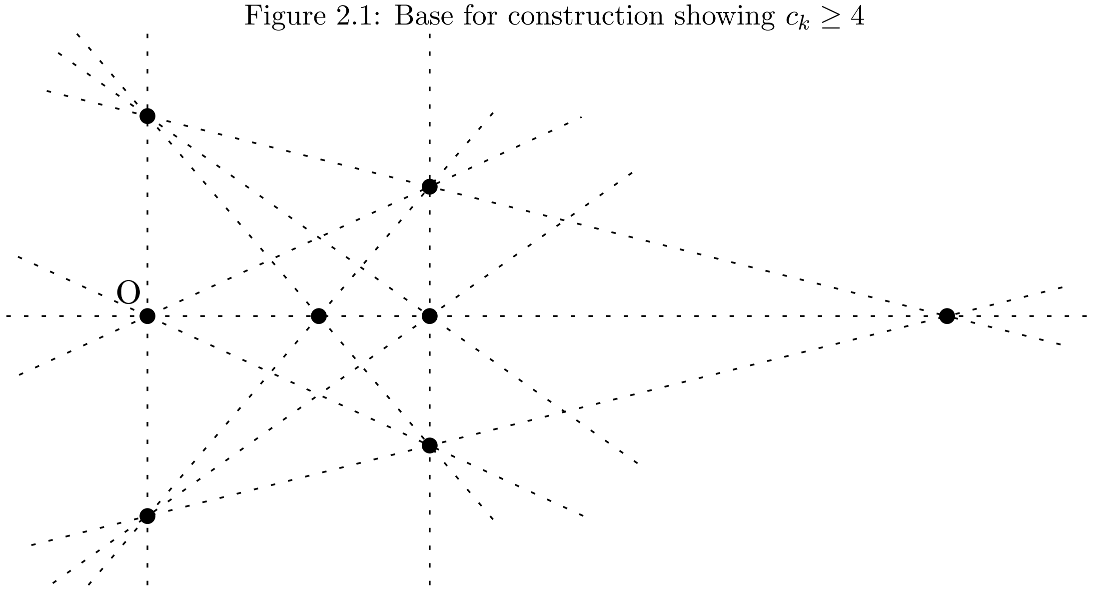
Grünbaum and Shephard found and catalogued simplicial arrangements of planes in real projective 3-space [33]. Among these are several examples that (after taking the dual arrangement of points) give sets of points that span more lines than planes, and that are not contained in a pair of lines, or in a plane and a point. In particular, the arrangement $A_1^3(18)$ gives a set of 18 points, spanning 60 planes and 74 lines, such that no plane or pair of lines contains more than 9 of the points. Later, Alexanderson and Wetzel [2] found an additional simplicial arrangement of planes. In the projective dual, this arrangement gives a set of 21 points, spanning 90 planes and 98 lines, such that no plane or pair of lines contains more than 10 of the points.

We can apply Lemma 26 with Alexanderson and Wetzel's construction. By taking a generic point as the origin, and $|L|$ sufficiently large, this construction gives $c_3 \geq 11$.

For $k = 2$, the hypercube example in section 2.7.2 gives the lower bound $c_2 \geq 2$. We now show a slightly more sophisticated construction that achieves the bound $c_2 \geq 4$.

Grünbaum has produced a lovely and useful catalog of the known simplicial line arrangements in the real projective plane [32]. We use one of the arrangements he describes as the foundation for the construction. In particular, the point set shown in figure 2.1 is dual to the arrangment $A(8, 1)$ in Grünbaum's catalog.

Figure 2.1: Base for construction showing $c_k \geq 4$



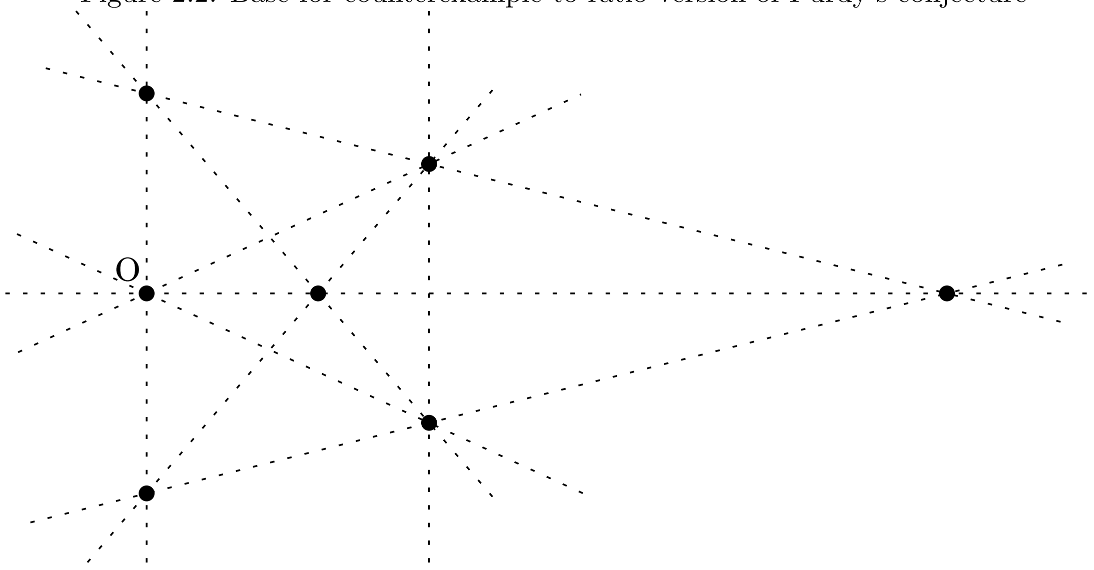
We apply Lemma 26 with the point set appearing in figure 2.1 as S, using the point marked "o" as the origin; i.e., let $P = S \cup L$, where S is the point set in figure 2.1, and L is a set of m collinear points contained in a line perpendicular to the plane spanned by S and incident to the point marked "o". By taking m to be sufficiently large, we can ensure that the points of L must be included in $\mathcal{G}_2$, and hence inspection of figure 2.1 shows that $n - g_2$ is 4. Further, we have $f_1^{\overline{o}}(S) = 7$, $f_1^{o}(S) = 4$, and $f_0^{\overline{o}}(S) = 7$.

Hence, Lemma 26 gives

$$f_2(P) = 7m + 4 + 1 + 0 < 7m + 1 + 11 = f_1(P),$$

and so this construction shows that $c_2 \geq 4$.

Figure 2.2: Base for counterexample to ratio version of Purdy's conjecture



In light of the preceeding examples, it might be tempting to conjecture that, under the hypothesis of Purdy's conjecture (i.e., P is a set of points that are not contained in the union of two lines or the union of a plane and a point), we at least have that $f_2 \geq f_1 - c$ for some universal constant c. However, even this weaker conjecture is too optimistic. To show this, we apply Lemma 26 with the point set appearing in figure 2.2 as S, using the point marked "o" as the origin. A brief examination of the figure reveals that $f_1^{\bar{o}}(S) = 5$ and $f_0^{\bar{o}}(S) = 6$, and that $n - g_2 = 3$. Hence, if we take m to be large, it follows from Lemma 26 that $f_2 < (5/6)f_1 + O(1)$.

2.8 A generalization of the Elekes-Tóth incidence bound

This section contains the proof of Theorem 13.

We first give a proof of the case $k = 3$. While this proof does not generalize to higher dimensions, it is very simple, and contains the germ of the idea of the full proof.

2.8.1 Proof of Theorem 13

2.8.2 A simple proof for the case $k = 3$

The case $k = 3$ admits a simpler proof than the general theorem.

Theorem 28. *For any $\alpha < 1$, the number of α-degenerate, r-rich 3-flats is bounded above by $O(n^4 r^{-5} + n^3 r^{-3})$.*

Proof. By Theorem 12, the number of essentially-$\alpha^{1/2}$-degenerate r-rich 3-flats is bounded above by $O(n^4 r^{-5} + n^3 r^{-3})$. If an r-rich 3-flat Λ is α-degenerate but not essentially-$\alpha^{1/2}$-degenerate, then at least $\alpha^{1/2}|P \cap \Lambda| \geq \alpha^{1/2} r$ points of P are contained in the union of two skew lines, neither of which contains more than $\alpha|P \cap \Lambda|$ points of P; hence, each of these lines contains at least $(\alpha^{1/2} - \alpha)r$ points. By the Szemerédi-Trotter theorem, the maximum number of pairs of $((\alpha^{1/2} - \alpha)r)$-rich lines is bounded above by $O(n^4 r^{-6} + n^2 r^{-2})$, which implies the conclusion of the theorem. $\square$

2.8.3 Proof of the general case

The proof of Theorem 28 given above does not generalize to higher dimensions, but the basic approach of bounding the number of r-rich α-degenerate flats that are not also essentially-α'-degenerate does still work in higher dimensions. This idea is captured by the following lemma.

Lemma 29. *Let $\mathcal{F} = \mathcal{F}_{\alpha,r}$ be the set of k-flats satisfying the following property. If $\Lambda \in \mathcal{F}$, then Λ contains a set $\mathcal{G}$ of flats so that*

1. *$\sum_{\Gamma \in \mathcal{G}} \dim(\Gamma) < k$,*
2. *$\overline{\mathcal{G}} = \Lambda$,*
3. *each flat of $\mathcal{G}$ is r-rich and α-degenerate.*

Then, $|\mathcal{F}| = O(n^{k+1} r^{-k-2} + n^k r^{-k})$.

We remark that Lemma 29 is not tight, in general; for example, a stronger bound of $O(n^4 r^{-6} + n^2 r^{-2})$ was given for the case $k = 3$ in the proof of Theorem 28, above.

Before proving Lemma 29, we show that it implies Theorem 13.

Proof of Theorem 13. Let $\alpha' < 1$ be a constant to fix later. The required bound on the number of r-rich, essentially-α'-degenerate k-flats is given by Theorem 12, so it only remains to bound the number of r-rich, α-degenerate k-flats that are not also essentially-α'-degenerate.

If Λ is an r-rich, α-degenerate k-flat that is not essentially-α'-degenerate, then there is a collection $\mathcal{G}'$ of flats with $\sum_{\Gamma \in \mathcal{G}'} \dim(\Gamma) < k$ such that $|\bigcup_{\Gamma \in \mathcal{G}'} \Gamma \cap P| > \alpha' |P \cap \Lambda|$.

We obtain a set $\mathcal{G}$ satisfying the conditions of Lemma 29 from $\mathcal{G}'$ as follows. If $\Gamma \in \mathcal{G}'$ is not α'-degenerate, then replace Γ with the smallest subspace $\Gamma' \subset \Gamma$ that contains at least $(\alpha')^{\dim(\Gamma) - \dim(\Gamma')} |P \cap \Gamma|$ points; note that Γ' is α'-degenerate, and we have removed fewer than $(\dim(\Gamma) - \dim(\Gamma'))(1 - \alpha')|P \cap \Lambda|$ points.

Next, remove from $\mathcal{G}'$ any flat that contains fewer than $(1 - \alpha')|P \cap \Lambda|$ points to obtain the final set $\mathcal{G}$. Each remaining flat in $\mathcal{G}$ is α'-degenerate and $(1-\alpha')|P \cap \Lambda|$-rich, and we have removed in all fewer than $\sum_{\Gamma \in \mathcal{G}'} \dim(\Gamma)(1 - \alpha')|P \cap \Lambda| < (\alpha' - \alpha)|P \cap \Lambda|$ points, for a sufficiently large choice of $\alpha' < 1$. If $\dim(\overline{\mathcal{G}}) < k$, then Λ is α-degenerate, contrary to our assumption. Hence, $\dim(\overline{\mathcal{G}}) = k$, and hence Λ belongs to the set $\mathcal{F}$ of Lemma 29. The conclusion of Lemma 29 implies the required bound, which completes the proof.

□

Here comes the proof of Lemma 29.

Proof of Lemma 29. We proceed by induction on k and $\sum_{\Gamma \in \mathcal{G}} \dim(\Gamma)$. The case $k = 2$ is trivial, and the case $k = 3$, $\sum_{\Gamma \in \mathcal{G}} \dim(\Gamma) = 2$ is Theorem 28.

We partition $\mathcal{F}$ into subsets $\mathcal{F}_b$, for $1 \leq b \leq k$, and separately bound the size of each $\mathcal{F}_b$.

Let $\Lambda \in \mathcal{F}$, and let $\mathcal{G}_\Lambda$ be the set of flats given by the hypothesis of Lemma 29. Let $\Gamma_\Lambda \in \mathcal{G}_\Lambda$, and let $b_\Lambda = \dim(\overline{\mathcal{G}_\Lambda \setminus \Gamma \Lambda})$. Assign Λ to $\mathcal{F}_{b_\Lambda}$.

With this assignment, the bound on $|\mathcal{F}_k|$ follows directly by the inductive hypothesis on $\sum_{\Gamma \in \mathcal{G}} \dim(\Gamma)$ by ignoring the flats Γ_Λ for $\Lambda \in \mathcal{F}_k$. For $b < k$, the inductive hypothesis implies that

$$|\{\overline{\mathcal{G}_\Lambda} : \Lambda \in \mathcal{F}_b\}| = O(n^{b+1} r^{-b-2} + n^b r^{-b}).$$

Hence, it will suffice to show that each $\mathcal{G}_\Lambda$ in this set is associated to at most $O(n^{k-b}r^{-k+b})$ different flats $\Lambda \in \mathcal{F}_b$.

Let $\mathcal{R} \in \{\overline{\mathcal{G}_\Lambda} : \Lambda \in \mathcal{F}_b\}$. Let $\pi_\mathcal{R}$ be the projection of $P \setminus (P \cap \mathcal{R})$ from $\mathcal{R}$; this is a multiset of points in $\mathbb{R}^{d-1-b}$, and if Λ is a flat such that $\dim(\overline{\Lambda, \mathcal{R}}) = k$, then $\dim(\pi_\mathcal{R})(\Lambda) = k - 1 - b$. Furthermore, if $\pi_\mathcal{R}(\Lambda)$ is α-degenerate, then so is Λ. To complete the proof, we will use the following lemma, proved below.

Lemma 30. *Let M be a multiset of points with total multiplicity n. The number of r-rich, α-degenerate k-flats spanned by M is bounded above by $(1-\alpha)^{-k}n^{k+1}r^{-k-1}$.*

From Lemma 30, we get the required bound of $O(n^{k-b}r^{b-k})$ on the number of r-rich, α-degenerate flats that can span a k-flat together with $\mathcal{R}$, and this completes the proof of Lemma 29. $\square$

Proof of Lemma 30. There are n^{k+1} ordered lists of $k+1$ points in M (with repetitions allowed). We show below that for any r-rich, α-degenerate k-flat Λ, the are at least $(1-\alpha)^k r^{k+1}$ distinct lists of $k+1$ points such that all of the points are contained in Λ, and the points are affinely independent. Since $k+1$ affinely independent points are contained in exactly one k-flat, an averaging argument completes the proof.

Let Λ be an r-rich, α-degenerate k-flat. We will show, by induction, that, for each $0 \le k' \le k$, Λ contains $(1-\alpha)^{k'}r^{k'+1}$ distinct ordered lists of $k'+1$ affinely independent points. The base case of $k' = 0$ is immediate from the fact that Λ is r-rich.

Choose uniformly at random a pair $(\mathbf{v}, p)$, where $\mathbf{v}$ is an ordered list of k' affinely independent points contained in Λ, and p is a point of P contained in Λ. By the inductive hypothesis, we know that there are $(1-\alpha)^{k'-1}r^{k'}$ choices for $\mathbf{v}$, and there are clearly $|P \cap \Lambda| \ge r$ choices for p. If the probability that p is affinely dependent on the points of $\mathbf{v}$ is more than α, then there is some $\mathbf{v}$ for which the number of points in $P \cap \Lambda$ that are affinely dependent on $\mathbf{v}$ is more than $\alpha|P \cap \Lambda|$. Since these points must all be contained in the $k'-1$-dimensional span of $\mathbf{v}$, this contradicts the hypothesis that Λ is α-degenerate. Hence, the number of choices of $(\mathbf{v}, p)$ such that p is affinely independent of $\mathbf{v}$ is at least $(1-\alpha)^{k'}r^{k'+1}$, which is what was to be proved. $\square$

Chapter 3

Distinct perpendicular bisectors and distances

3.1 Introduction

Many classic problems in discrete geometry ask for the minimum number of distinct equivalence classes of subsets of a fixed set of points under some geometrically defined equivalence relation. The seminal example is the Erdős distinct distance problem [25]: How few distinct distances can be determined by a set of n points in the Euclidean plane? Guth and Katz have nearly resolved the Erdős distinct distance question [35], but there are numerous other examples of questions of this type, many of which remain wide open.

One natural question that has not received much attention is: How few distinct perpendicular bisectors can be determined by a set of n points in the Euclidean plane? Distinct perpendicular bisectors were previously investigated by the author, Sheffer, and de Zeeuw [46], and a finite field analog was studied by Hanson, the author, and Roche-Newton [36].

Without any additional assumption, it is not too hard to give a complete answer to this question. The vertices of a regular n-gon determine n distinct perpendicular bisectors. Each point of an arbitrary point set $\mathcal{P}$ determines $n-1$ distinct bisectors with the remaining points of $\mathcal{P}$, and this is tight when $n=2$. In subsection 3.1.2, we give a simple geometric argument showing that the number of distinct bisectors is at least n, when $n>2$.

Suppose we assume that no circle or line contains more than K points of $\mathcal{P}$. In this case, the author, Sheffer, and de Zeeuw give the following lower bound on $|\mathcal{B}|$, the number of distinct bisectors determined by $\mathcal{P}$, a fixed set of n points in the Euclidean plane [46]:

$$|\mathcal{B}| = \Omega\left(\min\left\{K^{-\frac{2}{5}}n^{\frac{8}{5}-\varepsilon}, K^{-1}n^2\right\}\right). \tag{3.1}$$

We further proposed the following conjecture.

Conjecture 31. *For any $\varepsilon > 0$, there is a constant $c_\varepsilon > 0$ such that either a single line or circle contains $(1-\varepsilon)n$ points of P, or $|\mathcal{B}| \geq c_\varepsilon n^2$.*

In this chapter, we take a significant qualitatitive step toward Conjecture 31.

Theorem 32. *For any $\delta, \varepsilon > 0$, either a single circle or line contains $(1-\delta)n$ points of $\mathcal{P}$, or*

$$|\mathcal{B}| = \Omega(n^{52/35-\varepsilon}),$$

where the constants hidden in the Ω-notation depend on δ, ε.

This improves on the earlier result (3.1) of the author, Sheffer, and de Zeeuw for $K = \Omega(n^{2/7+\varepsilon})$, and gives the first non-trivial result on Conjecture 31 for $K = \Omega(n)$.

The proof (in [46]) of inequality (3.1) uses the, now standard, method of bounding the "energy"[1] of the quantity in question. In particular, we write $\mathcal{B}(a, b)$ for the perpendicular bisector of distinct points a, b, and define the *bisector energy* to be the size of the set

$$\mathcal{Q} = \{(a, b, c, d) \in \mathcal{P}^4 : a \neq b, c \neq d, \mathcal{B}(a, b) = \mathcal{B}(c, d)\}.$$

It is easy to see that $|\mathcal{Q}| \leq n^2(n-1)$, since each element of $\mathcal{Q}$ is determined by (a, b, c); taking $\mathcal{P}$ to be the vertices of a regular n-gon shows that this bound is tight. In [46], we show

$$|\mathcal{Q}| \leq O\left(K^{\frac{2}{5}} n^{\frac{12}{5}+\varepsilon} + Kn^2\right), \tag{3.2}$$

and conjecture that the strongest possible bound is $|\mathcal{Q}| \leq O(Kn^2)$. A standard application of Cauchy-Schwarz (see, for example, the proof of Lemma 39, below) gives

$$|\mathcal{B}| \geq n^2(n-1)^2/|\mathcal{Q}|.$$

Using this inequality, it is a straightforward calculation to obtain (3.1) from (3.2).

Observe that even a tight bound of $|\mathcal{Q}| \leq O(Kn^2)$ would give $|\mathcal{B}| \geq \Omega(n^2K^{-1})$. This only meets the bound of Conjecture 31 when K is a constant not depending on n, and does not give any non-trivial bound for $K = \Omega(n)$. Hence, it initially seems hopeless to use an energy bound to make substantial progress toward Conjecture 31 for

[1]The term *additive energy*, referring to the number of quadruples (a, b, c, d) in some underlying set of numbers such that $a+b = c+d$, was coined by Tao and Vu [37]. Starting with the work of Sharir, Elekes [22] and Guth, Katz [35] on the distinct distance problem, the strategy of using geometric incidence bounds to obtain upper bounds on analogously defined energies has become indispensable in the study of questions about the number of distinct equivalent subsets.

large K.

The main new idea in this chapter is to apply an energy bound to a refined subset of the pairs of points of $\mathcal{P}$. We show that there is a large set $\Pi \subset \mathcal{P} \times \mathcal{P}$ of pairs of points, such that

$$\mathcal{Q}^* = \{(a,b,c,d) \in \mathcal{P}^4 : (a,b),(c,d) \in \Pi, \mathcal{B}(a,b) = \mathcal{B}(c,d)\}$$

is small. In particular, we define Π to be the set of pairs of points of $\mathcal{P}$ that are not contained in any circle or line that contains too many points of $\mathcal{P}$. We use a point-circle incidence bound, proved in [5], to show that Π must be large, and use an argument similar to that bounding $\mathcal{Q}$ in [46] to show that $\mathcal{Q}^*$ must be small.

The proof of Theorem 32 is in Section 3.2.

3.1.1 Application to pinned distances

In Section 3.3, we give an application of the methods and results of this chapter to a problem of Erdős on the set of distances determined by a set of points in the plane. In particular, we give an alternate proof of a known bound.

Let P be a set of n points in the Euclidean plane. We denote

$$\delta(p) = \{\|x - p\| : x \in P\}$$
$$\delta^* = \max_{p \in P} |\delta(p)|$$

Erdős conjectured that $\delta^* = \Omega(n/\log(n))$ for all point sets. The best current result on this problem is by Katz and Tardos [41], who built on the work of Solymosi and Tóth [56]; Katz and Tardos showed that $\delta^* = \Omega(n^{0.864})$.

We give an alternate proof of the following, weaker, result.

Theorem 33. *For any $\varepsilon > 0$,*

$$\delta^* = \Omega(n^{87/105-\varepsilon}).$$

In the proof of Theorem 33, we use a new weighted Szemerédi-Trotter bound. This gives an upper bound on the number of incidences between weighted points and lines, when we have a bound on the sum of squares of the weights. Researchers working on similar problems in incidence geometry may find this a convenient tool. The statement and proof of the result are in Section 3.5.

3.1.2 There are at least n bisectors

We give the best possible general lower bound on $|\mathcal{B}|$.

Proposition 34. *If* $n > 2$*, then* $|\mathcal{B}| \geq n$.

Proof. Since any point $a \in \mathcal{P}$ determines $n-1$ distinct bisectors with the remaining points $\mathcal{P} \setminus \{a\}$, it is sufficient to show that there are three points a, b, c such that $\mathcal{B}(b, c)$ is distinct from $\mathcal{B}(a, x)$ for any $x \in \mathcal{P}$. If there are three collinear points, this is immediate, so we assume that no three points are collinear.

Let $a, b \in \mathcal{P}$ so that $|ab|$ is minimal, and let $c \in \mathcal{P}$ so that the angle $\angle abc$ is minimal. If a is on the same side of $\mathcal{B}(b, c)$ as c, then $|ac| \leq |ab|$, which is a contradiction. If a is on the line $\mathcal{B}(b, c)$, then there is no point x such that $\mathcal{B}(a, x) = \mathcal{B}(b, c)$, and we have accomplished our goal. Hence, we may suppose that a and b are on the same side of $\mathcal{B}(b, c)$. Let x be the reflection of a over $\mathcal{B}(b, c)$. The line ax is parallel to the line bc, and x and c are on the same side of ab. Hence, x is in the interior of the cone defined by $\angle abc$, and hence $\angle abx$ is less than $\angle abc$. Since c was chosen so that $\angle abc$ is minimal, $x \notin \mathcal{P}$, which completes the proof. $\square$

3.1.3 Acknowledgements

I would like to thank Brandon Hanson, Peter Hajnal, Oliver Roche-Newton, Adam Sheffer, and Frank de Zeeuw for many stimulating conversations on perpendicular bisectors and related questions. I would also like to thank Luca Ghidelli for pointing out an error in Lemma 38 in an earlier version.

3.2 Proof of Theorem 32

In this section, we prove Theorem 32.

Handling heavy circles.

We first apply a separate, elementary argument to handle the case that a single circle contains a substantial portion of the points of $\mathcal{P}$.

Lemma 35. *If a single line or circle contains exactly εn points of $\mathcal{P}$, then*

$$|\mathcal{B}| \geq \min(\varepsilon, 1-\varepsilon) \cdot \varepsilon n^2/4.$$

We rely on the following geometric lemma.

Lemma 36. *Let C be a circle or a line, and let $p, q \notin C$ with $p \neq q$. Then,*

$$\#\{(r,s) \in C \times C : \mathcal{B}(p,r) = \mathcal{B}(q,s)\} \leq 2.$$

Proof. Fix $p, q \notin C$. For $r \in C$, let C_r be the reflection of C over $\mathcal{B}(p,r)$; note that $C_r = C_{r'}$ implies $r = r'$. If $s \in C$ such that $\mathcal{B}(p,r) = \mathcal{B}(q,s)$, then $q \in C_r$. Since there are two circles that are the same size as C and that contain p and q, there are at most two pairs $(r,s) \in C \times C$ such that $\mathcal{B}(p,r) = \mathcal{B}(q,s)$. $\square$

Proof of Lemma 35. Let C be a circle that contains εn points of $\mathcal{P}$. Let $\mathcal{P}' \subset \mathcal{P}$ be a set of $k = \min(\varepsilon, 1-\varepsilon)n$ points that are not in C. Let $p_1, p_2, \ldots p_k$ be an arbitrary ordering of the points of $\mathcal{P}'$. Then, by Lemma 36, p_i determines a set $\mathcal{B}(p_i)$ of at least $\varepsilon n - 2(i-1)$ distinct perpendicular bisectors with the points of P that lie on C, such that no element of $\mathcal{B}(p_i)$ is an element of $\mathcal{B}(p_j)$ for any $j < i$. Summing over i, we have

$$\sum_{i \leq k} |\mathcal{B}(p_i)| \geq \varepsilon n k/4,$$

which proves the lemma. $\square$

Lemma 35 implies that, if the maximum number of points of $\mathcal{P}$ that are contained in any circle is at least cn (for a constant c to be determined later) and at most $(1 - O(n^{-18/35+\epsilon})n = (1-\delta)n$ points, then $|\mathcal{B}| = \Omega(n^{52/35-\varepsilon})$. Hence, we may assume from now on that no circle contains more than cn points.

Refining the pairs of points.

Now we handle the case that no circle contains more than some small, constant fraction of the points of $\mathcal{P}$.

Denote by s_k the number of lines and circles that contain at least k points of $\mathcal{P}$, denote by $s_{=k}$ the number of lines and circles that contain exactly k points of $\mathcal{P}$.

The following incidence bound combines the result of Szemerédi and Trotter [58] on point-line incidences with the result of Aronov and Sharir [5] on point-circle incidences.

Lemma 37 (Combined point-line and point-circle incidence bound)**.** *For any $\varepsilon > 0$,*

$$s_k = O(n^{3+\varepsilon}k^{-11/2} + n^2k^{-3} + nk^{-1}),$$

where the hidden constants depend on ε.

We apply Lemma 37 to show that either a single circle contains a constant fraction of the points of $\mathcal{P}$, or a constant fraction of the pairs of points in $\mathcal{P}\times\mathcal{P}$ are not contained in any line or circle that contains too many points of $\mathcal{P}$.

Lemma 38. *For any $\varepsilon > 0$, there are constants $c_1, c_2 > 0$ such that the following holds. Let $\mathcal{P}$ be a set of n points. Let $\Pi \subset \mathcal{P} \times \mathcal{P}$ be the set of pairs of distinct points of $\mathcal{P}$ such that no pair in Π is contained in a line or a circle that contains $M = c_1 n^{2/7+\varepsilon}$ points of $\mathcal{P}$. Then, either there is a single line or circle that contains $c_2 n$ points of $\mathcal{P}$, or $|\Pi| = \Omega(n^2)$.*

Proof. By Lemma 37, the number of triples (p, q, C) of two points p, q and a line or a circle C such that $p, q \in C$ and such that C contains at least M and at most $U = c_2 n$

points is bounded above by

$$\begin{aligned}
\sum_{k=M}^{U} k^2 s_{=k} &= \sum_{k=M}^{U} k^2 (s_k - s_{k+1}) \\
&\leq \sum_{k=M}^{U} 2k s_k \\
&\leq O\left(\sum_{k \geq M} n^{3+\varepsilon} k^{-9/2} + \sum_{k \geq M} n^2 k^{-2} + \sum_{k \leq U} n \right) \\
&\leq O\left(n^2\right).
\end{aligned}$$

With appropriate choices of c_1, c_2, we can ensure that the constant hidden in the O-notation on the final line is less than 1. □

Let $\Pi \subseteq \mathcal{P} \times \mathcal{P}$ be the set of pairs of distinct points of $\mathcal{P}$ that do not lie on any circle that contains more than $M = c_1 n^{2/7+\varepsilon}$ points of $\mathcal{P}$. We have already handled the case that a single line or circle contains $c_2 n$ points of $\mathcal{P}$, and so we assume that this does not occur, and consequently (by Lemma 38) that $|\Pi| = \Omega(n^2)$.

Let

$$\begin{aligned}
&\mathcal{B}^* = \{\mathcal{B}(a,b) : (a,b) \in \Pi\}, \text{ and} \\
&\mathcal{Q}^* = \{(a,b,c,d) : (a,b),(c,d) \in \Pi, \mathcal{B}(a,b) = \mathcal{B}(c,d)\}.
\end{aligned}$$

An application of Cauchy-Schwarz produces a lower bound on $|\mathcal{B}^*|$ (and hence, on $|\mathcal{B}|$) from an upper bound on $|\mathcal{Q}^*|$.

Lemma 39.

$$|\mathcal{B}| = \Omega\left(n^4 |\mathcal{Q}^*|^{-1}\right).$$

Proof. For a line ℓ, denote by $w(\ell)$ the number of pairs $(p,q) \in \Pi$ such that $\mathcal{B}(p,q) = \ell$.

By Cauchy-Schwarz,

$$\begin{aligned}
|\mathcal{Q}^*| &= \sum_{\ell \in \mathcal{B}^*} w(\ell)^2, \\
&\geq \left(\sum_{\ell \in \mathcal{B}^*} w(\ell) \right)^2 |\mathcal{B}^*|^{-1}, \\
&= |\Pi|^2 |\mathcal{B}^*|^{-1}.
\end{aligned}$$

Hence,

$$\begin{aligned}
|\mathcal{B}| &\geq |\mathcal{B}^*|, \\
&\geq |\Pi|^2 |\mathcal{Q}^*|^{-1}, \\
&= \Omega\left(n^4 |\mathcal{Q}^*|^{-1}\right).
\end{aligned}$$

□

Bounding the energy.

We will use another incidence geometry argument to bound $|Q^*|$; this part of the analysis has substantial overlap with the proof of Theorem 2.1 in [46].

For each pair $(a, b) \in P^2$, let $C(a, b)$ be the maximum number of points on any circle that contains a, b. Let

$$\begin{aligned}
\Pi_K &= \{(a, b) \in P^2 : a \neq b, C(a, b) \leq K\}, \\
\mathcal{Q}_K &= \{(a, b, c, d) \in P^4 : (a, b), (c, d) \in \Pi_K, \mathcal{B}(a, b) = \mathcal{B}(c, d)\}.
\end{aligned}$$

We prove

Lemma 40. *For any $2 \leq K \leq n$,*

$$|Q_K| = O\left(K^{\frac{2}{5}} n^{\frac{12}{5}+\varepsilon} + K n^2\right).$$

Proof. For each pair (a, c) of distinct points in $\mathcal{P}$, we define the *bisector surface* to be

$$S_{ac} = \{(b, d) \in \mathbb{R}^2 \times \mathbb{R}^2 : \ \mathcal{B}(a, b) = \mathcal{B}(c, d)\},$$

and we define

$$\mathcal{S} = \{S_{ac} : a, c \in \mathcal{P}, a \neq c\}.$$

The following is [46, Lemma 3.1].

Lemma 41. *For distinct $a, c \in \mathcal{P}$, there exists a two-dimensional constant-degree algebraic variety $\overline{S}_{ac}$ such that $S_{ac} \subset \overline{S}_{ac}$. Moreover, if $(b, d) \in (\overline{S}_{ac} \backslash S_{ac})$ with $b \neq d$, then either $a = b$ or $c = d$.*

We denote

$$\overline{\mathcal{S}} = \{\overline{S}_{ac} : a, c \in \mathcal{P}, a \neq c\}.$$

Let $G \subset \overline{\mathcal{S}} \times \mathcal{P}^2$ be the incidence graph between pairs of distinct points of $\mathcal{P}$ and varieties in $\overline{\mathcal{S}}$. Let $H \subset \mathcal{S} \times \mathcal{P}^2$ be the incidence graph between pairs of distinct points of $\mathcal{P}$ and surfaces in $\mathcal{S}$. Let $G' \subset \mathcal{S} \times \mathcal{P}^2$ such that $(S_{ac}, (b, d)) \in G'$ if and only if $(b, d) \in S_{ac}$ and $(a, b), (c, d) \in \Pi_K$. By identifying the vertices corresponding to S_{ac} and $\overline{S}_{ac}$ for each a, c, we have $G' \subseteq H \subseteq G$.

Note that

$$G' = \{(S_{ac}, (b, d)) \in \mathcal{S} \times \mathcal{P}^2 : (a, b), (c, d) \in \Pi_K, B(a, c) = B(b, d)\},$$

and hence

$$|G'| = |\mathcal{Q}^*|.$$

Observe that if $\mathcal{B}(a, b) = \mathcal{B}(c, d)$, then the reflection of the pair (a, c) over the line $\mathcal{B}(a, b)$ is the pair (b, d). Hence, $|ac| = |bd|$. It follows that, if $|ac| = \delta$, then the surface S_{ac} is contained in the hypersurface

$$H_\delta = \{(b, d) \in \mathbb{R}^2 \times \mathbb{R}^2 : |bd| = \delta\}.$$

The following is [46, Lemma 3.2].

Lemma 42. *Let $a, c \in \mathbb{R}^2, (a,c) \neq (a',c')$ and $|ac| = |a'c'| = \delta \neq 0$. Then there exist curves $C_1, C_2 \subset \mathbb{R}^2$, which are either two concentric circles or two parallel lines, such that $a, a' \in C_1$, $c, c' \in C_2$, and $S_{ac} \cap S_{a'c'}$ is contained in the set*

$$H_\delta \cap (C_1 \times C_2) = \{(b,d) \in \mathbb{R}^2 \times \mathbb{R}^2 : b \in C_1, d \in C_2, |bd| = \delta\}.$$

We use Lemmas 41 and 42 to prove

Lemma 43. *If $\overline{S}_{ac}, \overline{S}_{a'c'}$ have $K+4$ or more common neighbors in G, then $S_{ac}, S_{a'c'}$ have no common neighbors in G'.*

Proof. We claim that, if S_{ab} and $S_{a'b'}$ share K or more common neighbors in H, then they have no common neighbors in G'. Suppose that S_{ac} and $S_{a'c'}$ share K common neighbors in H; denote this set of pairs of points $(b,d) \in S_{ac} \cap S_{a'c'}$ by Γ. Lemma 42 implies that there exist two lines or circles C_1, C_2 with $a, a' \in C_1$ and $c, c' \in C_2$ such that $(b,d) \in \Gamma$ only if $b \in C_1$ and $d \in C_2$, and $|bd| = \delta$. Note that $(b,d), (b,d') \in S_{ac}$ implies that $\mathcal{B}(c,d) = \mathcal{B}(a,b) = \mathcal{B}(c,d')$, and hence $d = d'$. Hence, there are at least K distinct points in each of C_1 and C_2, and so, if $(b,d) \in \Gamma$, then $(a,b), (a',b), (c,d), (c',d) \notin \Pi_K$. Hence, if $(b,d) \in \Gamma$, then $(S_{ac}, (b,d))$ and $(S_{a'c'}, (b,d))$ are not in G'.

Suppose that $a \neq a'$ and $c \neq c'$. By Lemma 41, each neighbor of $\overline{S}_{ac}$ in $G \setminus H$ is either of the form (a,d) or (b,c). Hence, a vertex (b,d) can be a common neighbor of S_{ac} and $S_{a'c'}$ in G but not in H only if $b = a$ or $b = a'$ or $d = c$ or $d = c'$. Since S_{ac} is incident to at most one point (b,x) for any fixed b, this implies that $S_{ac}, S_{a'c'}$ have at most 4 common neighbors in G that are not common neighbors in H. Hence, if $\overline{S}_{ac}$ and $\overline{S}_{a'c'}$ have $K+4$ or more common neighbors in G, then S_{ac} and $S_{a'c'}$ have K or more common neighbors in H, and by the previous claim have no common neighbors in G'.

Now, suppose that $a = a'$ (the case $c = c'$ is symmetric). Since $(b,d) \in S_{ac} \cap S_{ac'}$ implies that $c = c'$, we have that S_{ac} and $S_{ac'}$ have no common neighbors in H. Hence, they have no common neighbors in G'.

$\square$

Let $\delta_1, \ldots, \delta_D$ denote the distinct non-zero distances determined by pairs of distinct points in $\mathcal{P}$. Let

$$\mathcal{P}_i^2 = \{(b,d) \in \mathcal{P} \times \mathcal{P} : |ab| = \delta_i\},$$
$$\mathcal{S}_i = \{S_{ac} \in \mathcal{S} : |ac| = \delta_i\},$$
$$G_i' = \{(S_{ac},(b,d)) \in G' : |pq| = \delta_i\}.$$

Let

$$m_i = |\mathcal{P}_i^2| = |\mathcal{S}_i|.$$

As observed above, each quadruple $(a,b,c,d) \in Q$ satisfies $|ac| = |bd|$. Hence, it suffices to study each G_i' separately. That is, we have

$$|Q_K| = |G'| = \sum_{i=1}^{D} |G_i'|.$$

We will use the following incidence bound to control the size of each $|G_i'|$. This bound is a slight generalization of a bound in [46], which is in turn a generalization of a bound in [29]. See [29] for definitions of the algebraic terms used.

Theorem 44. *Let $\mathcal{S}$ be a set of n constant-degree varieties, and let $\mathcal{P}$ be a set of m points, both in $\mathbb{R}^d$. Let $s \geq 2$ be a constant, and $t \geq 2$ be a function of m, n. Let G be the incidence graph of $\mathcal{P} \times \mathcal{S}$. Let $G' \subseteq G$ such that, if a set L of s left vertices has a common neighborhood of size t or more in G, then no pair of vertices in L has a common neighbor in G'. Moreover, suppose that $\mathcal{P} \subset V$, where V is an irreducible constant-degree variety of dimension e. Then*

$$|G'| = O\left(m^{\frac{s(e-1)}{es-1}+\varepsilon} n^{\frac{e(s-1)}{es-1}} t^{\frac{e-1}{es-1}} + tm + n\right).$$

The proof of Theorem 44 is nearly identical to the proof of Theorem 2.5 in [46], requiring only one small, technical change. For the sake of completeness, the proof of

Theorem 44 is included in Appendix 3.6.

We apply Theorem 44 to the set of varieties $\overline{\mathcal{S}}_i = \{\overline{S}_{ac} : S_{ac} \in \mathcal{S}_i\}$, the set of points $\mathcal{P}_i^2$, and G and G' as the corresponding incidence graphs. The hypersurface H_{δ_i} is irreducible, three-dimensional, and of constant degree, since it is defined by the irreducible polynomial $(x_1 - x_3)^2 + (x_2 - x_4)^2 - \delta_i$. Thus, we can apply Theorem 44 with $m = n = m_i$, $V = H_{\delta_i}$, $d = 4$, $e = 3$, $s = 2$, and $t = K$. Hence,

$$|G_i'| = O(K^{2/5} m_i^{7/5+\varepsilon} + K m_i). \tag{3.3}$$

Let J be the set of indexes $1 \le j \le D$ for which the bound in (3.3) is dominated by the term $K^{\frac{2}{5}} m_j^{\frac{7}{5}+\varepsilon}$. By recalling that $\sum_{j=1}^{D} m_j = n(n-1)$, we get

$$\sum_{j \notin J} |G_j'| = O\left(Kn^2\right).$$

Next we consider $\sum_{j\in J} |G_j'| = O(\sum_{j \in J} K^{2/5} m_j^{7/5+\varepsilon})$. By [35, Proposition 2.2], we have

$$\sum m_j^2 = O(n^3 \log n).$$

This implies that the number of m_j for which $m_j \ge x$ is $O(n^3 \log n / x^2)$. Using a dyadic decomposition, we obtain

$$\begin{aligned} K^{-2/5} n^{-\varepsilon} \sum_{j \in J} |G_j'| &= O\left(\sum_{m_j \le \Delta} m_j^{7/5} + \sum_{k \ge 1} \sum_{2^{k-1}\Delta < m_j \le 2^k \Delta} m_j^{7/5} \right) \\ &= O\left(\Delta^{7/5} \cdot \frac{n^2}{\Delta} + \sum_{k \ge 1} (2^k \Delta)^{\frac{7}{5}} \cdot \frac{n^3 \log n}{(2^{k-1}\Delta)^2} \right) \\ &= O\left(\Delta^{2/5} n^2 + \frac{n^3 \log n}{\Delta^{3/5}} \right). \end{aligned}$$

By setting $\Delta = n \log n$, we have

$$\sum_{j \in J} |G_j'| = O\left(K^{\frac{2}{5}} n^{\frac{12}{5}+\varepsilon} \log^{\frac{2}{5}} n \right) = O\left(K^{\frac{2}{5}} n^{\frac{12}{5}+\varepsilon'} \right).$$

Combining the bounds, we have

$$|\mathcal{Q}^*| = |G'| = \sum_{i=1}^{D} |G'_i| = O\left(K^{\frac{2}{5}} n^{\frac{12}{5}+\varepsilon'} + Kn^2\right),$$

which completes the proof of Lemma 40. □

Finishing the proof. Taking $K = M = c_1 n^{2/7+\varepsilon}$ in Lemma 40, we have

$$|\mathcal{Q}^*| = O(n^{88/35+\varepsilon}).$$

Combining this with Lemma 39, we have

$$|\mathcal{B}| = \Omega(n^{52/35-\varepsilon}),$$

which is Theorem 32.

3.3 Proof of Theorem 33

We prove Theorem 33 by double counting the set

$$\Delta = \{(a, b, c) \in P^3 : |ab| = |ac|, b \neq c\},$$

which is the set of oriented, non-degenerate isoceles triangles determined by P.

The lower bound on Δ proceeds by a standard application of the Cauchy-Schwarz inequality. We denote by $n(p, \delta)$ the number of points of P at distance δ from p.

$$\begin{aligned}
|\Delta| &= \sum_{p \in P} \sum_{\delta \in \delta(p)} (n(p, \delta) - 1)^2, \\
&\geq \sum_{p \in P} (n - |\delta(p)|)^2 |\delta(p)|^{-1}, \\
&\geq \sum_{p \in P} (n - \delta^*)^2 (\delta^*)^{-1}, \\
&= n(n - \delta^*)^2 (\delta^*)^{-1}.
\end{aligned}$$

For the upper bound, we apply the observation that the number of isosceles triangles is equal to the number of incidences between the points of P and bisectors of P, counted with multiplicity. We denote the multiplicity of a bisector $\ell \in \mathcal{B}$ by

$$w(\ell) = |\{(a,b) \in P^2 : \mathcal{B}(a,b) = \ell\}|.$$

It is easy to see that

$$|\Delta| = I(P,\mathcal{B}) = \sum_{p\in P}\sum_{\ell\in\mathcal{B}}[p \in \ell]w(\ell).$$

The notation $[p \in \ell]$ denotes the indicator function that takes value 1 if $p \in \ell$ and 0 otherwise.

Recall that, for each pair $(a,b) \in P^2$, we denote by $C(a,b)$ the maximum number of points of P on any circle that contains both a and b. Let

$$\begin{aligned}
\Pi_{k,K} &= \{(a,b) \in P^2 : a \neq b, k \leq C(a,b) < K\},\\
\mathcal{B}_{k,K} &= \{\mathcal{B}(a,b) : (a,b) \in \Pi_{k,K}\},\\
w_{k,K}(\ell) &= |\{(a,b) \in \Pi_{k,K} : \mathcal{B}(a,b) = \ell\}|.
\end{aligned}$$

We decompose the incidences as follows:

$$I(P,\mathcal{B}) = I(P,\mathcal{B}_{2,M}) + \sum_{\log M \leq i < \log n} I(P,\mathcal{B}_{2^i,2^{i+1}}), \tag{3.4}$$

in which incidences between P and $\mathcal{B}_{k,K}$ are weighted by $w_{k,K}$, and, as in Section 3.2, $M = c_1 n^{2/7+\varepsilon}$.

We will use Theorem 48, proved in Section 3.5, to bound the size of those sets of incidences in (3.4) involving bisector mulitiplicities at most $n^{1/2}$. Applying Theorem 37, we have

$$\sum_{\ell\in\mathcal{B}_{k,K}} w(\ell) = |\Pi_{k,K}| = \min(n^2, O(k^2(n^{3+\varepsilon}k^{-11/2} + n^2k^{-3} + nk^{-1}))). \tag{3.5}$$

Applying Lemma 40, we have

$$\sum_{\ell \in \mathcal{B}_{k,K}} w(\ell)^2 = |Q_{k,K}| = O(K^{2/5} n^{12/5+\varepsilon} + Kn^2). \tag{3.6}$$

It is clear that no line can be the bisector of more than n pairs of points.

Recall that $Q^* = Q_{2,M}$ and $\Pi = \Pi_{2,M}$. Applying Theorem 48 together with (3.5) and (3.6), we have

$$I(P, \mathcal{B}_{2,M}) = O(n^{2/3}|Q^*|^{1/3}|\Pi|^{1/3} + n^2) = O(n^{228/105+\varepsilon}). \tag{3.7}$$

Dividing the remaining range for $2^i < n^{1/2}$ depending on which terms in (3.5) and (3.6) are dominant, straightforward calculations show that

$$I(P, \mathcal{B}_{k,2k}) = O(n^{37/15+\varepsilon'} k^{-31/30}), \qquad M \le k < c_a n^{2/5+\varepsilon}, \tag{3.8}$$

$$I(P, \mathcal{B}_{k,2k}) = O(n^{32/15+\varepsilon'} k^{-1/5}), \qquad c_a n^{2/5+\varepsilon} \le k < c_b n^{1/2}. \tag{3.9}$$

For $k \ge c_b n^{1/2}$, note that Theorem 37 implies that there are $O(nk^{-1})$ circles that each contain at least k points. Let $\mathcal{C}$ be the set of circles that contain between k and $2k$ points, for some $c_b n^{1/2} < k < n$. Let p be an arbitrary point of P, and let C be an arbitrary circle in $\mathcal{C}$. If p is the center of C, then p is incident to all bisectors determined by pairs of points on C, which have total multiplicity $O(k^2)$. Since $|\mathcal{C}| = O(nk^{-1})$, there are at most so many centers of circles in $\mathcal{C}$, so the total number of such incidences is $O(nk) = O(n^2)$. Otherwise, p is incident to at most one bisector determined by pairs of points on C, which has total multiplicity $O(k)$. Since $|\mathcal{C}| = O(nk^{-1})$, there are $O(n)$ such incidences between p and circles of C, and so the total number of such incidences is $O(n^2)$.

Hence, for $c_b n^{1/2} \le k \le n/2$, we have

$$I(P, \mathcal{B}_{k,2k}) = O(n^2). \tag{3.10}$$

Note that the right sides of (3.8), (3.9), and (3.10) are all bounded above by

$O(n^{228/105+\varepsilon})$. Hence, substituting into (3.4), we have

$$I(P, \mathcal{B}) = O(n^{228/105+\varepsilon} \log(n)).$$

Absorbing the $\log(n)$ into n^{ε}, this gives us the upper bound

$$|\Delta| = O(n^{228/105+\varepsilon}).$$

Combining the upper and lower bounds, we have

$$n(n - \delta*)^2(\delta^*)^{-1} \leq |\Delta| = O(n^{228/105+\varepsilon}),$$

which implies

$$\delta^* = \Omega(n^{87/105-\varepsilon}).$$

3.4 Discussion

The proofs of Theorems 32 and 33 both depend on Theorem 37 and Lemma 40, and neither of these are tight. Any improvement in the bounds for Theorem 37 or Lemma 40 will immediately translate to corresponding improvements to Theorems 32 and 33. Indeed, if the following conjectured bounds are proved for Theorem 37 and Lemma 40, we will immediately have a nearly tight bound for perpendicular bisectors in place of Theorem 32.

A tight bound on the number of incidences between points and circles would imply a nearly tight bound for Theorem 33 with no need for an improvement to Lemma 40; simply place a circle around each point $p \in \mathcal{P}$ for each distance from p, and then directly bound the number of point-circle incidences. Note that a tight bound for refined bisector energy together with Theorem 37 would not immediately imply a tight bound for the Erdős pinned distance problem, but would improve the bound of Katz and Tardos [41].

Let P be a set of n points in the Euclidean plane. Recall that c_k denotes the maximum number of circles that contain at least k points of P.

Conjecture 45. *For any $\varepsilon > 0$ and $k > n^{\varepsilon}$,*

$$c_k = O(n^2 k^{-3} + nk^{-1}).$$

Recall that $C(a, b)$ denotes the maximum number of points on any circle that contains $a, b \in P^2$, and

$$\Pi_K = \{(a, b) \in P^2 : a \neq b, C(a, b) \leq K\},$$
$$\mathcal{Q}_K = \{(a, b, c, d) \in P^4 : (a, b), (c, d) \in \Pi_K, \mathcal{B}(a, b) = \mathcal{B}(c, d)\}.$$

Conjecture 46.

$$|\mathcal{Q}_K| = O(Kn^2).$$

The proofs of Theorems 32 and 33 in Sections 3.2 and 3.3 can easily be adapted to use Conjectures 45 and 46. If Conjectures 45 and 46 were proved, we would immediately have (for any $\varepsilon > 0$) the bound

$$|\mathcal{B}| = \Omega(n^{2-\varepsilon}), \tag{3.11}$$

assuming that no more than cn points lie on any circle, for some $c < 1$. Both of these bounds would be tight up to the n^{ε} factors.

3.5 Weighted Szemerédi-Trotter

In this section, we prove a generalized Szemerédi-Trotter theorem for weighted points and lines, which is useful when we have control over the sum-of-squares of the weights. This strengthens an earlier weighted Szemerédi-Trotter from [38], in the case that the weights of lines or points differ substantially.

For a set A with a weight function $w : A \to \mathbb{Z}^+$, let

$$|A|_1 = \sum_{a \in A} w(a),$$
$$|A|_2^2 = \sum_{a \in A} w(a)^2,$$
$$|A|_\infty = \max_{a \in A} w(a).$$

For a weighted set P of points and a weighted set L of lines, define

$$I(P, L) = \sum_{p \in P} \sum_{\ell \in L} [p \in \ell] w(p) w(l)$$

to be the number of weighted incidences between P and L.

We need the standard Szemerédi-Trotter theorem for unweighted points and lines.

Theorem 47 (Szemerédi-Trotter)**.** *Let P be a set of points, and L a set of lines, in $\mathbb{R}^2$. Then,*

$$I(P, L) = O(|P|^{2/3}|L|^{2/3} + |P| + |L|).$$

Theorem 48 (Weighted Szemerédi-Trotter)**.** *Let P be a set of weighted points, and L a set of weighted lines, in $\mathbb{R}^2$. Then,*

$$I(P, L) = O\left((|P|_2^2 |P|_1 |L|_2^2 |L|_1)^{1/3} + |L|_\infty |P|_1 + |P|_\infty |L|_1 \right).$$

Proof. Let

$$L_i = \{\ell \in L : 2^i \le w(\ell) < 2^{i+1}\},$$
$$P_i = \{p \in P : 2^i \le w(p) < 2^{i+1}\}.$$

Then, applying a dyadic decomposition and Theorem 47, we have

$$\begin{aligned} I(P,L) &= \sum_{\ell\in L}\sum_{p\in P}[p\in\ell]w(p)w(\ell), \\ &\ll \sum_{1\le 2^i<|L|_\infty}\ \sum_{1\le 2^j<|P|_\infty} 2^{i+1}2^{j+1}(|L_i|^{2/3}|P_j|^{2/3}+|L_i|+|P_j|), \end{aligned} \tag{3.12}$$

Since

$$|L|_1=\sum_{\ell\in L}w(\ell)\ge \sum_{1\le 2^i<|L|_\infty}2^i|L_i|,$$

we have

$$\sum_{1\le 2^i<|L|_\infty}\ \sum_{1\le 2^j<|P|_\infty}2^i2^j|L_i|\le |L|_1\sum_{1\le 2^j<|P|_\infty}2^j\ll |L|_1|P|_\infty. \tag{3.13}$$

Similarly,

$$\sum_{1\le 2^i<|L|_\infty}\ \sum_{1\le 2^j<|P|_\infty}2^i2^j|P_i|\ll |P|_1|L|_\infty. \tag{3.14}$$

Next, we bound the term $\sum_{1\le 2^i<|L|_\infty}2^i|L_i|^{2/3}$ in (3.12). We split the sum as

$$\sum_{1\le 2^i<|L|_\infty}2^i|L_i|^{2/3}=\sum_{1\le 2^i<|L|_2^2|L|_1^{-1}}2^i|L_i|^{2/3}+\sum_{|L|_2^2|L|_1^{-1}\le 2^i<|L|_\infty}2^i|L_i|^{2/3}.$$

Note that $\sum_i 2^i|L_i|\le |L|_1$, and hence $2^i|L_i|\le |L|_1$ for any particular i. Also note that $\sum_{1\le 2^i<|L|_2^2|L|_1^{-1}}2^{i/3}$ is only a constant factor larger than its largest term, $|L|_2^{2/3}|L|_1^{-1/3}$.

$$\begin{aligned} \sum_{1\le 2^i<|L|_2^2|L|_1^{-1}}2^i|L_i|^{2/3} &\le \sum_{1\le 2^i<|L|_2^2|L|_1^{-1}}2^i(|L|_12^{-1})^{2/3}, \\ &= |L|_1^{2/3}\sum_{1\le 2^i<|L|_2^2|L|_1^{-1}}2^{i/3}, \\ &\ll |L|_1^{1/3}|L|_2^{2/3}. \end{aligned}$$

Note that $\sum_i 2^{2i}|L_i| \leq |L|_2^2$, and hence $2^{2i}|L_i| \leq |L|_2^2$ for any particular i.

$$\begin{aligned}\sum_{|L|_2^2|L|_1^{-1}\leq 2^i<|L|_\infty} 2^i|L_i|^{2/3} &\leq \sum_{|L|_2^2|L|_1^{-1}\leq 2^i<|L|_\infty} 2^i(|L|_2^2 2^{-2i})^{2/3},\\ &= |L|_2^{4/3} \sum_{|L|_2^2|L|_1^{-1}\leq 2^i<|L|_\infty} 2^{-i/3},\\ &\ll |L|_1^{1/3}|L|_2^{2/3}.\end{aligned}$$

Hence,

$$\sum_{1\leq 2^i<|L|_\infty} |L_i|^{2/3} \ll |L|_1^{1/3}|L|_2^{2/3}, \tag{3.15}$$

and similarly,

$$\sum_{1\leq 2^j<|P|_\infty} |P_j|^{2/3} \ll |P|_1^{1/3}|L|_2^{2/3}. \tag{3.16}$$

Combining (3.13), (3.14), (3.15), and (3.16) with (3.12) completes the proof. □

3.6 Proof of Theorem 44

The proof of Theorem 44 is nearly identical to the proof of Theorem 2.5 in [46]. The main difference occurs in bounding the quantity $|I_1|$ (defined below).

The proof uses the Kővári-Sós-Turán theorem (see for example [9, Theorem IV.9]).

Lemma 49 (Kővári-Sós-Turán)**.** *Let G be a bipartite graph with vertex set $A \cup B$. Let $s \leq t$. Suppose that G contains no $K_{s,t}$; that is, for any s vertices in A, at most $t-1$ vertices in B are connected to each of the s vertices. Then*

$$|G| = O(t^{\frac{1}{s}}|A||B|^{\frac{s-1}{s}} + |B|).$$

We amplify the weak bound of Lemma 49 by using *polynomial partitioning.* Given a polynomial $f \in \mathbb{R}[x_1,\dots,x_d]$, we write $Z(f) = \{p \in \mathbb{R}^d : f(p) = 0\}$. We say that $f \in \mathbb{R}[x_1,\dots,x_d]$ is an *r-partitioning polynomial* for a finite set $\mathcal{P} \subset \mathbb{R}^d$ if no connected component of $\mathbb{R}^d\backslash Z(f)$ contains more than $|\mathcal{P}|/r$ points of $\mathcal{P}$ (notice that there is no restriction on the number of points of $\mathcal{P}$ that are in $Z(f)$). Guth and Katz [35] introduced this notion and proved that for every $\mathcal{P} \subset \mathbb{R}^d$ and $1 \leq r \leq |\mathcal{P}|$, there exists an r-partitioning polynomial of degree $O(r^{1/d})$. In [29], the following generalization was proved.

Theorem 50 (Partitioning on a variety)**.** *Let V be an irreducible variety in $\mathbb{R}^d$ of dimension e and degree D. Then for every finite $\mathcal{P} \subset V$ there exists an r-partitioning polynomial f of degree $O(r^{1/e})$ such that $V \not\subset Z(f)$. The implicit constant depends only on d and D.*

We are now ready to prove the incidence bound.

Proof of Theorem 44. Note that we may assume that no variety in $\mathcal{S}$ contains V. We can assume that V contains at least s points (otherwise the bound in the theorem is trivial). If there are at most $t-1$ varieties in $\mathcal{S}$ that contain V, then these varieties altogether give less than tm incidences, which is accounted for in the bound. If there are t or more varieties in $\mathcal{S}$ that countain V, then Lemma 43 implies that no pair of vertices in G' corresponding to a pair of points contained in V shares any neighbor among the vertices corresponding to the varieties that contain V. These are at most m incidences, which is accounted for in the bound.

We use induction on e and m, with the induction claim being that for $\mathcal{P}, \mathcal{S}, V, G'$ as in the theorem, with the added condition that no variety in $\mathcal{S}$ contains V, we have

$$|G'| \leq \alpha_{1,e} m^{\frac{s(e-1)}{es-1}+\varepsilon} n^{\frac{e(s-1)}{es-1}} t^{\frac{e-1}{es-1}} + \alpha_{2,e}(tm+n), \tag{3.17}$$

for constants $\alpha_{1,e}, \alpha_{2,e}$ depending only on d, e, s, ε, the degree of V, and the degrees of the varieties in $\mathcal{S}$. The base cases for the induction are simple. If m is sufficiently small, then (3.17) follows immediately by choosing sufficiently large values for $\alpha_{1,e}$ and $\alpha_{2,e}$. Similarly, when $e = 0$, we again obtain (3.17) when $\alpha_{1,e}$ and $\alpha_{2,e}$ are sufficiently large (as a function of d and the degree of V).

The constants d, e, s, ε are given and thus fixed, as are the degree of V and the degrees of the varieties in $\mathcal{S}$. The other constants are to be chosen, and the dependencies between them are

$$C_{\text{weak}}, C_{\text{part}}, C_{\text{inter}} \ll C_{\text{cells}} \ll C_{\text{Höld}} \ll r \ll C_{\text{comps}} \ll \alpha_{2,e} \ll \alpha_{1,e},$$

where $C \ll C'$ means that C' is to be chosen sufficiently large compared to C; in particular, C should be chosen before C'. Furthermore, the constants $\alpha_{1,e}, \alpha_{2,e}$ depend on $\alpha_{1,e-1}, \alpha_{2,e-1}$.

Note that G' is $K_{s,t}$-free. Hence, by Lemma 49, there exists a constant C_{weak} depending on d, s such that

$$|G'| \leq C_{\text{weak}} \left(mn^{1-\frac{1}{s}} t^{\frac{1}{s}} + n \right).$$

When $m \leq (n/t)^{1/s}$, and $\alpha_{2,e}$ is sufficiently large, we have $|G'| \leq \alpha_{2,e} n$. Therefore, in the remainder of the proof we can assume that $n < m^s t$, which implies

$$n = n^{\frac{e-1}{es-1}} n^{\frac{e(s-1)}{es-1}} \leq m^{\frac{s(e-1)}{es-1}} n^{\frac{e(s-1)}{es-1}} t^{\frac{(e-1)}{es-1}}. \tag{3.18}$$

Partitioning.

By Theorem 50, there exists an r-partitioning polynomial f with respect to V of degree at most $C_{\text{part}} \cdot r^{1/e}$, for a constant C_{part}. Denote the cells of $V \backslash Z(f)$ as $\Omega_1, \ldots, \Omega_N$. Since we are working over the reals, there exists a constant-degree polynomial g such that $Z(g) = V$. Then, by [55, Theorem A.2], the number of cells is bounded by $C \cdot \deg(f)^{\dim V} = C_{\text{cells}} \cdot r$, for some constant C_{cells} depending on C_{part}.

We partition G' into the following three subsets:

- I_1 consists of the incidences $(p, S) \in \mathcal{P} \times \mathcal{S}$ such that $p \in V \cap Z(f)$, and some irreducible component of $V \cap Z(f)$ contains p and is fully contained in S.
- I_2 consists of the incidences $(p, S) \in \mathcal{P} \times \mathcal{S}$ such that $p \in V \cap Z(f)$, and no irreducible component of $V \cap Z(f)$ that contains p is contained in S.
- $I_3 = G' \backslash (I_1 \cup I_2)$, the set of incidences $(p, S) \in \mathcal{P} \times \mathcal{S}$ such that p is not contained in $V \cap Z(f)$.

Note that we indeed have $G' = I_1 \cup I_2 \cup I_3$, excluding any edges corresponding to varieties in $\mathcal{S}$ that fully contain V.

Bounding $|I_1|$.

The points of $\mathcal{P} \subset \mathbb{R}^d$ that participate in incidences of I_1 are all contained in the variety $V_0 = V \cap Z(f)$. Set $\mathcal{P}_0 = \mathcal{P} \cap V_0$ and $m_0 = |\mathcal{P}_0|$. Since V is an irreducible variety and $V \not\subset Z(f)$, V_0 is a variety of dimension at most $e-1$ and of degree that depends on r. By [55, Lemma 4.3], the intersection V_0 is a union of C_{comps} irreducible components, where C_{comps} is a constant depending on r and d.[2] The degrees of these components also depend only on these values (for a proper definition of degrees and further discussion, see for instance [29]).

Consider an irreducible component W of V_0. If W contains at most $s-1$ points of $\mathcal{P}_0$, it yields at most $(s-1)n$ incidences. Otherwise, if there are at most $t-1$ varieties of $\mathcal{S}$ that fully contain W, then these yield at most $(t-1)m_0$ incidences. Otherwise, if there are at least t varieties of $\mathcal{S}$ that fully contain W, then by Lemma 43, no pair of vertices corresponding to points contained in W has a common neighbor in G' among the varieties that contain W. In this case, at most m_0 incidences must be counted.

By summing up, choosing sufficiently large $\alpha_{1,e}, \alpha_{2,e}$, and applying (3.18), we have

$$|I_1| \leq C_{\text{comps}}\left(sn + tm_0\right) < \frac{\alpha_{2,e}}{2}(n+tm_0) < \frac{\alpha_{1,e}}{4} m^{\frac{s(e-1)}{es-1}} n^{\frac{e(s-1)}{es-1}} t^{\frac{(e-1)}{es-1}} + \frac{\alpha_{2,e}}{2} tm_0. \quad (3.19)$$

Bounding $|I_2|$.

The points that participate in I_2 lie in $V_0 = V \cap Z(f)$, and the varieties that participate do not contain any component of V_0. Because V_0 has dimension at most $e-1$, we can apply the induction claim on each irreducible component W of V_0, for the point set $\mathcal{P} \cap W$ and the set of varieties in $\mathcal{S}$ that do not contain W. Since V_0 has C_{comps}

---

[2] This lemma only applies to complex varieties. However, we can take the *complexification* of the real variety and apply the lemma to it (for the definition of a complexification, see for example [66, Section 10]). The number of irreducible components of the complexification cannot be smaller than number of irreducible components of the real variety (see for instance [66, Lemma 7]).

irreducible components, we get

$$|I_2| \le C_{\text{comps}} \alpha_{1,e-1} m_0^{\frac{s(e-2)}{(e-1)s-1}+\varepsilon} n^{\frac{(e-1)(s-1)}{(e-1)s-1}} t^{\frac{e-2}{(e-1)s-1}} + \alpha_{2,e-1}(tm_0 + n),$$

with $\alpha_{1,e-1}$ and $\alpha_{2,e-1}$ depending on the degree of the irreducible component of V_0, which in turn depends on r. Recalling that we may assume $n < m^s t$, we obtain

$$\begin{aligned} m^{\frac{s(e-2)}{(e-1)s-1}+\varepsilon} n^{\frac{(e-1)(s-1)}{(e-1)s-1}} t^{\frac{e-2}{(e-1)s-1}} &= m^{\frac{s(e-2)}{(e-1)s-1}+\varepsilon} n^{\frac{e(s-1)}{es-1}} n^{\frac{s-1}{(es-s-1)(es-1)}} t^{\frac{e-2}{(e-1)s-1}} \\ &< m^{\frac{s(e-1)}{es-1}+\varepsilon} n^{\frac{e(s-1)}{es-1}} t^{\frac{e-1}{es-1}}. \end{aligned}$$

By applying (3.18) to remove the term $\alpha_{2,e-1}n$, and by choosing $\alpha_{1,e}$ and $\alpha_{2,e}$ sufficiently large as a function of $C_{\text{comps}}, \alpha_{1,e-1}, \alpha_{2,e-1}$, we obtain

$$|I_2| \le \frac{\alpha_{1,e}}{4} m^{\frac{s(e-1)}{es-1}+\varepsilon} n^{\frac{e(s-1)}{es-1}} t^{\frac{e-1}{es-1}} + \frac{\alpha_{2,e}}{2} tm_0. \tag{3.20}$$

Bounding $|I_3|$.

For every $1 \le i \le N$, we set $\mathcal{P}_i = \mathcal{P} \cap \Omega_i$ and denote by $\mathcal{S}_i$ the set of varieties of $\mathcal{S}$ that intersect the cell Ω_i. Let $G_i' \subseteq G'$ be $(\mathcal{P}_i \times \mathcal{S}_i) \cap G'$. We also set $m_i = |P_i|$ and $n_i = |\mathcal{S}_i|$. Then we have $m_i \le m/r$ and $\sum_{i=1}^N m_i = m - m_0$.

Let $S \in \mathcal{S}$. By the assumption made at the beginning of the proof, S does not contain V, so $S \cap V$ is a subvariety of V of dimension at most $e-1$. By [55, Theorem A.2], there exists a constant C_{inter} such that the number of cells intersected by $S \cap V$ is at most $C \cdot \deg(f)^{\dim(S \cap V)} = C_{\text{inter}} \cdot r^{(e-1)/e}$. This implies that

$$\sum_{i=1}^N n_i \le C_{\text{inter}} \cdot r^{\frac{e-1}{e}} \cdot n.$$

By Hölder's inequality we have

$$\begin{aligned}\sum_{i=1}^{N} n_i^{\frac{e(s-1)}{es-1}} &\le \left(\sum_{i=1}^{N} n_i\right)^{\frac{e(s-1)}{es-1}} \left(\sum_{i=1}^{N} 1\right)^{\frac{e-1}{es-1}}\\ &\le \left(C_{\text{inter}} r^{(e-1)/e} n\right)^{\frac{e(s-1)}{es-1}} (C_{\text{cells}} r)^{\frac{e-1}{es-1}}\\ &\le C_{\text{Höld}} r^{\frac{(e-1)s}{es-1}} n^{\frac{e(s-1)}{es-1}},\end{aligned}$$

where $C_{\text{Höld}}$ depends on $C_{\text{inter}}, C_{\text{cells}}$. Using the induction claim for each i with the point set $\mathcal{P}_i$, the set of varieties $\mathcal{S}_i$, and the same variety V, we obtain

$$\begin{aligned}\sum_{i=1}^{N} |G_i'| &\le \sum_{i=1}^{N} \left(\alpha_{1,e} m_i^{\frac{(e-1)s}{es-1}+\varepsilon} n_i^{\frac{e(s-1)}{es-1}} t^{\frac{(e-1)}{es-1}} + \alpha_{2,e}(tm_i + n_i)\right)\\ &\le \alpha_{1,e} \frac{m^{\frac{(e-1)s}{es-1}+\varepsilon} t^{\frac{(e-1)}{es-1}}}{r^{\frac{(e-1)s}{es-1}+\varepsilon}} \sum_{i=1}^{N} n_i^{\frac{e(s-1)}{es-1}} + \sum_{i=1}^{N} \alpha_{2,e}(tm_i + n_i)\\ &\le \alpha_{1,e} C_{\text{Höld}} \frac{m^{\frac{(e-1)s}{es-1}+\varepsilon} n^{\frac{e(s-1)}{es-1}} t^{\frac{(e-1)}{es-1}}}{r^{\varepsilon}} + \alpha_{2,e}\left(t(m-m_0) + C_{\text{inter}} r^{\frac{e-1}{e}} n\right).\end{aligned}$$

By choosing $\alpha_{1,e}$ sufficiently large with respect to $C_{\text{inter}}, r, \alpha_{2,e}$, and using (3.18), we get

$$\sum_{i=1}^{N} |G_i'| \le 2\alpha_{1,e} C_{\text{Höld}} \frac{m^{\frac{(e-1)s}{es-1}+\varepsilon} n^{\frac{e(s-1)}{es-1}} s^{\frac{(e-1)}{es-1}}}{r^{\varepsilon}} + \alpha_{2,e} t(m - m_0).$$

Finally, choosing r sufficiently large with respect to $C_{\text{Höld}}$ gives

$$|I_3| = \sum_{i=1}^{N} I(\mathcal{P}_i, \mathcal{S}_i) \le \frac{\alpha_{1,e}}{2} m^{\frac{(e-1)s}{es-1}+\varepsilon} n^{\frac{e(s-1)}{es-1}} t^{\frac{(e-1)}{es-1}} + \alpha_{2,e} t(m - m_0). \tag{3.21}$$

Summing up.

By combining $|G'| = |I_1| + |I_2| + |I_3|$ with (3.19), (3.20), and (3.21), we obtain

$$|G'| \le \alpha_{1,e} m^{\frac{s(e-1)}{es-1}+\varepsilon} n^{\frac{e(s-1)}{es-1}} t^{\frac{(e-1)}{es-1}} + \alpha_{2,e}(tm + n),$$

which completes the induction step and the proof of the theorem. □

Chapter 4

Incidence bounds for block designs

4.1 Introduction

The structure of incidences between points and various geometric objects is of central importance in discrete geometry, and theorems that elucidate this structure have had applications to, for example, problems from discrete and computational geometry [35, 49], additive combinatorics [13, 21], harmonic analysis [34, 42], and computer science [20]. The study of incidence theorems for finite geometry is an active area of research - e.g. [64, 11, 15, 39, 40, 60, 24].

The classical Szemerédi-Trotter theorem [59] bounds the maximum number of incidences between points and lines in 2-dimensional Euclidean space. Let P be a set of points in $\mathbb{R}^2$ and L be a set of lines in $\mathbb{R}^2$. Let $I(P, L)$ denote the number of incidences between points in P and lines in L. The Szemerédi-Trotter Theorem shows that $I(P, L) = O(|P|^{2/3}|L|^{2/3} + |P| + |L|)$. Ever since the original result, variations and generalizations of such incidence bounds have been intensively studied.

Incidence theorems for points and flats[1] in finite geometries is one instance of such incidence theorems that have received much attention, but in general we still do not fully understand the behavior of the bounds in this setting. These bounds have different characteristics depending on the number of points and flats. For example, consider the following question of proving an analog to the Szemerédi-Trotter theorem for points and lines in a plane over a finite field. Let $q = p^n$ for prime p. Let P be a set of points and L a set of lines in $\mathbb{F}_q^2$, with $|P| = |L| = N$. What is the maximum possible value of $I(P, L)$ over all point sets of size N and sets of lines of size N?

If $N \leq O(\log\log\log(p))$, then a result of Grosu [31] implies that we can embed P and L in $\mathbb{C}^2$ without changing the underlying incidence structure. Then we apply the result from the complex plane, proved by Tóth [61] and Zahl [67], that $I(P, L) \leq O(N^{4/3})$. This matches the bound of Szemerédi and Trotter, and a well-known construction based on a grid of points in $\mathbb{R}^2$ shows that the exponent of 4/3 is tight.

The intermediate case of $N < p$ is rather poorly understood. A result of Bourgain, Katz, and Tao [11], later improved by Jones [40], shows that $I(P, L) \leq O(N^{3/2-\epsilon})$ for

[1]We refer to m-dimensional affine subspaces of a vector space as m-flats.

$\epsilon = 1/662 - o_p(1)$. This result relies on methods from additive combinatorics, and is far from tight; in fact, we are not currently aware of any construction with $N < p^{3/2}$ that achieves $I(P, L) > \omega(N^{4/3})$.

For $N \geq q$, we know tight bounds on $I(P, L)$. Using an argument based on spectral graph theory, Vinh [64] proved that $|I(P, L) - N^2/q| \leq q^{1/2}N$, which gives both upper and lower bounds on $I(P, L)$. The upper bound meets the Szemerédi-Trotter bound of $O(N^{4/3})$ when $N = q^{3/2}$, which is tight by the same construction used in the real plane. The lower bound becomes trivial for $N \leq q^{3/2}$, and Vinh showed [65] that there are sets of $q^{3/2}$ points and $q^{3/2}$ lines such that there are no incidences between the points and lines. When $N = q$, we have $I(P, L) = O(N^{3/2})$, which is tight when $q = p^2$ for a prime power p (consider incidences between all points and lines in $\mathbb{F}_p^2$).

In this chapter we generalize Vinh's argument to the purely combinatorial setting of balanced incomplete block designs (BIBDs). This shows that his argument depends only on the combinatorial structure induced by flats in the finite vector space. We apply methods from spectral graph theory. We both generalize known incidence bounds for points and flats in finite geometries, and prove results for BIBDs that are new even in the special case of points and flats in finite geometries. Finally, we apply one of these incidence bounds to improve a result of Iosevich, Rudnev, and Zhai [39] on the number of triangles with distinct areas determined by a set of points in $\mathbb{F}_q^2$.

4.1.1 Outline

In Section 4.2.1, we state definitions of and basic facts about BIBDs and finite geometries. In Section 4.2.2, we discuss our results on the incidence structure of designs. In Section 4.2.3, we discuss our result on distinct triangle areas in $\mathbb{F}_q^2$. In Section 4.3, we introduce the tools from spectral graph theory that are used to prove our incidence results. In Section 4.4, we prove the results stated in Section 4.2.2. In Section 4.5, we prove the result stated in Section 4.2.3.

4.2 Results

4.2.1 Definitions and Background

Let X be a finite set (which we call the points), and let B be a set of subsets of X (which we call the blocks). We say that (X, B) is an (r, k, λ)-BIBD if

- each point is in r blocks,
- each block contains k points,
- each pair of points is contained in λ blocks, and
- no single block contains all of the points.

It is easy to see that the following relations among the parameters $|X|, |B|, r, k, \lambda$ of a BIBD hold:

$$\begin{aligned} r|X| &= k|B|, && (4.1) \\ \lambda(|X|-1) &= r(k-1). && (4.2) \end{aligned}$$

The first of these follows from double counting the pairs $(x, b) \in X \times B$ such that $x \in b$. The second follows from fixing an element $x \in X$, and double counting the pairs $(x', b) \in (X \setminus \{x\}) \times B$ such that $x', x \in b$.

In the case where X is the set of all points in $\mathbb{F}_q^n$, and B is the set of m-flats in $\mathbb{F}_q^n$, we obtain a design with the following parameters [6]:

- $|X| = q^n$,
- $|B| = \binom{n+1}{m+1}_q - \binom{n}{m+1}_q$,
- $r = \binom{n}{m}_q$,
- $k = q^m$, and
- $\lambda = \binom{n-1}{m-1}_q$.

The notation $\binom{n}{m}_q$ refers to the q-binomial coefficient, defined for integers $m \leq n$ by

$$\binom{n}{m}_q = \frac{(q^n-1)(q^n-q)\dots(q^n-q^{m-1})}{(q^m-1)(q^m-q)\dots(q^m-q^{m-1})}.$$

We will only use the fact that $\binom{n}{m}_q = (1+o_q(1))q^{m(n-m)}$.

Given a design (X, B), we say that a point $x \in X$ is incident to a block $b \in B$ if $x \in b$. For subsets $P \subseteq X$ and $L \subseteq B$, we define $I(P, L)$ to be the number of incidences between P and L; in other words,

$$I(P, L) = |\{(x, b) \in P \times L : x \in b\}|.$$

Given a subset $L \subseteq B$, we say that a point x is t-rich if it is contained in at least t blocks of L, and we define $\Gamma_t(L)$ to be the number of t-rich points in X; in other words,

$$\Gamma_t(L) = |\{x \in X : |\{b \in L : x \in b\}| \geq t\}|.$$

Given a subset $P \subseteq X$, we say that a block b is t-rich if it contains at least t points of P, and we define $\Gamma_t(P)$ to be the number of t-rich blocks in B; in other words,

$$\Gamma_t(P) = |\{b \in B : |\{x \in P : x \in b\}| \geq t\}|.$$

4.2.2 Incidence Theorems

The first result on the incidence structure of designs is a generalization of the finite field analog to the Szemerédi-Trotter theorem proved by Vinh [64].

Theorem 51. *Let (X, B) be an (r, k, λ)-BIBD. The number of incidences between $P \subseteq X$ and $L \subseteq B$ satisfies*

$$\big|\, I(P, L) - |P||L|r/|B| \,\big| \leq \sqrt{(r - \lambda)|P||L|}.$$

Theorem 51 gives both upper and lower bounds on the number of incidences between arbitrary sets of points and blocks. The term $|P||L|r/|B|$ corresponds to the number of incidences that we would expect to see between P and L if they were chosen uniformly at random. If $|P||L|$ is much larger than $|B|^2(r-\lambda)/r^2 > |B|^2/r$, then $|P||L|r/|B|$ is much larger than $\sqrt{(r - \lambda)|P||L|}$. Thus the theorem says that every set of points and blocks determines approximately the "expected" number of incidences. When $|P||L| < |B|^2/r$, the term on the right is larger, and Theorem 51 gives only an upper bound on the number of incidences. The Cauchy-Schwartz inequality combined with the fact that each pair of points is in at most λ blocks easily implies that $I(P,L) \leq \lambda^{1/2}|P||L|^{1/2}+|L|$. Hence, the upper bound in Theorem 51 is only interesting when $|P| > (r-\lambda)/\lambda$.

In the case of incidences between points and m-flats in $\mathbb{F}_q^n$, we get the following result as a special case of Theorem 51.

Corollary 52. *Let P be a set of points and let L be a set of m-flats in $\mathbb{F}_q^n$. Then*

$$\left|I(P,L) - |P|\,|L|q^{m-n}\right| \leq (1+o_q(1))\sqrt{q^{m(n-m)}|P|\,|L|}.$$

Vinh [64] proved Corollary 52 in the case $m = n-1$, and Bennett, Iosevich, and Pakianathan [8] derived the bounds for the remaining values of m from Vinh's bound, using elementary combinatorial arguments. Vinh's proof is based on spectral graph theory, analogous to the proof of Theorem 51 that we present in Section 4.4. Cilleruelo proved a result similar to Vinh's using Sidon sets [14].

We also show lower bounds on the number of t-rich blocks determined by a set of points, and on the number of t-rich points determined by a set of blocks.

While reading the statements of these theorems, it is helpful to recall from equation (1) that $|X|/k = |B|/r$.

Theorem 53. *Let (X,B) be an (r,k,λ)-BIBD. Let $\epsilon \in \mathbb{R}_{>0}$ and $t \in \mathbb{Z}_{\geq 2}$. Let $P \subseteq X$ with*

$$|P| \geq (1+\epsilon)(t-1)|X|/k.$$

Then, the number of t-rich blocks is at least

$$\Gamma_t(P) \geq a_{\epsilon,t,\mathcal{D}}|B|,$$

where

$$a_{\epsilon,t,\mathcal{D}} = \frac{\epsilon^2(t-1)}{\epsilon^2(t-1)+(1-\frac{\lambda}{r})(1+\epsilon)}.$$

Theorem 54. *Let (X,B) be an (r,k,λ)-BIBD. Let $\epsilon \in \mathbb{R}_{>0}$ and $t \in \mathbb{Z}_{\geq 2}$. Let $L \subseteq B$,* *with*

$$|L| \geq (1+\epsilon)(t-1)|X|/k.$$

Then, the number of t-rich points is at least

$$\Gamma_t(L) \geq b_{\epsilon,t,\mathcal{D}}|X|,$$

where

$$b_{\epsilon,t,\mathcal{D}} = \frac{\epsilon^2(t-1)}{\epsilon^2(t-1) + \frac{r-\lambda}{k}(1+\epsilon)}.$$

Theorem 53 is analogous to Beck's theorem [7], which states that, if P is a set of points in $\mathbb{R}^2$, then either $c_1|P|$ points lie on a single line, or there are $c_2|P|^2$ lines each contain at least 2 points of P, where c_1 and c_2 are fixed positive constants.

Both the $t = 2$ case of Theorem 53 and Beck's theorem are closely related to the de Bruijn-Erdős theorem [17, 54], which states that, if P is a set, and L is a set of subsets of P such that each pair of elements in P is contained in exactly λ members of L, then either a single member of L contains all elements of P, or $|L| \geq |P|$. Each of Theorem 53 and Beck's theorem has an additional hypothesis on the de Bruijn-Erdős theorem, and a stronger conclusion. Beck's theorem improves the de Bruijn-Erdős theorem when P is a set of points in $\mathbb{R}^2$ and L is the set of lines that each contain 2 points of P. Theorem 53 improves the de Bruijn-Erdős theorem when P is a sufficiently large subset of the points of a BIBD, and L is the set of blocks that each contain at least 2 points of P.

As special cases of Theorems 53 and 54, we get the following results on the number of t-rich points determined by a set of m-flats in $\mathbb{F}_q^n$, and on the number of t-rich m-flats determined by a set of points in $\mathbb{F}_q^2$.

Corollary 55. *Let $\epsilon \in \mathbb{R}_{\geq 0}$ and $t \in \mathbb{Z}_{\geq 2}$. Let $P \subseteq \mathbb{F}_q^n$ with*

$$|P| \geq (1+\epsilon)(t-1)q^{n-m}.$$

Then the number of t-rich m-flats is at least

$$\Gamma_t(L) \geq a_{\epsilon,t} q^{(m+1)(n-m)},$$

where

$$a_{\epsilon,t} = \frac{\epsilon^2(t-1)}{\epsilon^2(t-1) + (1+\epsilon)}.$$

Corollary 56. *Let $\epsilon \in \mathbb{R}_{\geq 0}$ and $t \in \mathbb{Z}_{\geq 2}$. Let L be a subset of the m-flats in $\mathbb{F}_q^n$ with*

$$|L| \geq (1+\epsilon)(t-1)q^{n-m}.$$

Then, the number of t-rich points is at least

$$\Gamma_t(L) \geq b_{\epsilon,t,q} q^n,$$

where

$$b_{\epsilon,t,q} = \frac{\epsilon^2(t-1)}{\epsilon^2(t-1) + q^{m(n-m-1)}(1+\epsilon)}.$$

The case $n = 2, m = 1, t = 2$ of Corollary 55 was proved (for a slightly smaller value of $a_{\epsilon,t}$) by Alon [3]. Alon's proof is also based on spectral graph theory.

When $m = n-1$, Corollaries 55 and 56 are dual (in the projective sense) to each other; slight differences in the parameters arise since we're working in affine (as opposed to projective) geometry.

For the case $m < n-1$, the value of $b_{\epsilon,t,q}$ in Corollary 56 depends strongly on q when $\epsilon(t-1) < q^{m(n-m-1)}$. This dependence is necessary. For example, consider the case $n = 3, m = 1$, i.e. lines in $\mathbb{F}_q^3$. Corollary 56 implies that a set of $2q^2$ lines in $\mathbb{F}_q^3$ determines $\Omega(q^2)$ 2-rich points, which is asymptotically fewer (with regard to q) than the total number of points in $\mathbb{F}_q^3$. This is tight, since the lines may lie in the union of two planes. By contrast, Corollary 55 implies that a set of $2q^2$ points in $\mathbb{F}_q^3$ determines $\Omega(q^4)$ 2-rich lines, which is a constant proportion of all lines in the space.

4.2.3 Distinct Triangle Areas

Iosevich, Rudnev, and Zhai [39] studied a problem on distinct triangle areas in $\mathbb{F}_q^2$. This is a finite field analog to a question that is well-studied in discrete geometry over the reals. Erdős, Purdy, and Strauss [27] conjectured that a set of n points in the real plane determines at least $\lfloor \frac{n-1}{2} \rfloor$ distinct triangle areas. Pinchasi [51] proved that this is the case.

In $\mathbb{F}_q^2$, we define the area of a triangle in terms of the determinant of a matrix. Suppose a triangle has vertices a, b, and c, and let z_x and z_y denote the x and y coordinates of a point z. Then, we define the area associated to the ordered triple (a, b, c) to be the determinant of the following matrix:

$$\begin{bmatrix} 1 & 1 & 1 \\ a_x & b_x & c_x \\ a_y & b_y & c_y \end{bmatrix}$$

Iosevich, Rudnev, and Zhai [39] showed that a set of at least $64q \log_2 q$ points includes a point that is a common vertex of triangles having at least $q/2$ distinct areas. They first prove a finite field analog of Beck's theorem, and then obtain their result on distinct triangle areas using this analog to Beck's theorem along with some Fourier analytic and combinatorial techniques. Corollary 55 (in the case $n = 2, m = 1, t = 2$) strengthens their analog to Beck's theorem, and thus we are able to obtain the following strengthening of their result on distinct triangle areas.

Theorem 57. *Let $\epsilon \in \mathbb{R}_{>0}$. Let P be a set of at least $(1+\epsilon)q$ points in $\mathbb{F}_q^2$. Let T be the set of triangles determined by P. Then there is a point $z \in P$ such that z is a common vertex of triangles in T with at least $c_\epsilon q$ distinct areas, where c_ϵ is a positive constant depending only on ϵ, such that $c_\epsilon \to 1$ as $\epsilon \to \infty$.*

Notice that Theorem 57 is tight in the sense that fewer than q points might determine only triangles with area zero (if all points are collinear). It is a very interesting open question to determine the minimum number of points K_q, such that any set of points of size K_q determines all triangle areas. In fact, we are not currently aware of any set of more than $q + 1$ points that does not determine all triangle areas.

4.3 Tools from Spectral Graph Theory

4.3.1 Context and Notation

Let $G = (L, R, E)$ be a (Δ_L, Δ_R)-biregular bipartite graph; in other words, G is a bipartite graph with left vertices L, right vertices R, and edge set E, such that each left vertex has degree Δ_L, and each right vertex has degree Δ_R. Let A be the $|L \cup R| \times |L \cup R|$ adjacency matrix of G, and let $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_{|L|+|R|}$ be the eigenvalues of A. Let $\mu = \mu_2/\mu_1$ be the normalized second eigenvalue of G.

Let $e(G) = \Delta_L|L| = \Delta_R|R|$ be the number of edges in G. For any two subsets of vertices A and B, denote by $e(A, B)$ the number of edges between A and B. For a subset of vertices $A \subseteq L \cup R$, denote by $\Gamma_t(A)$ the set of vertices in G that have at least t neighbors in A.

4.3.2 Lemmas

We will use two lemmas relating the normalized second eigenvalue of G to its combinatorial properties. The first of these is the expander mixing lemma [4].

Lemma 58 (Expander Mixing Lemma)**.** *Let $S \subseteq L$ with $|S| = \alpha|L|$ and let $T \subseteq R$ with $|T| = \beta|R|$. Then,*

$$\left| \frac{e(S, T)}{e(G)} - \alpha\beta \right| \leq \mu\sqrt{\alpha\beta(1-\alpha)(1-\beta)}.$$

Several variants of this result appear in the literature, most frequently without the $\sqrt{(1-\alpha)(1-\beta)}$ terms. For a proof that includes these terms, see [63], Lemma 4.15. Although the statement in [63] is not specialized for bipartite graphs, it is easy to modify it to obtain Lemma 58. For completeness, we include a proof in the appendix.

Theorems 54 and 53 follow from the following corollary to the expander mixing lemma, which may be of independent interest.

Lemma 59. *Let $\epsilon \in \mathbb{R}_{>0}$, and let $t \in \mathbb{Z}_{\geq 2}$. If $S \subseteq L$ such that*

$$|S| \geq (1+\epsilon)(t-1)|L|/\Delta_R,$$

then

$$|\Gamma_t(S)| \geq c_{\epsilon,t,G}|R|,$$

where

$$c_{\epsilon,t,G} = \frac{\epsilon^2(t-1)}{\epsilon^2(t-1) + \mu^2\Delta_R(1+\epsilon)}.$$

Proof. Let $T = R \setminus \Gamma_t(S)$. Let $\alpha = |S|/|L|$, and let $\beta = |T|/|R|$. We will calculate a lower bound on $1-\beta = |\Gamma_t(S)|/|R|$, from which we will immediately obtain a lower bound on $|\Gamma_t(S)|$.

Since each vertex in T has at most $t-1$ edges to vertices in S, we have $e(S,T) \leq (t-1)|T|$. Along with the fact that $|T| = \beta|R|$, this gives

$$\begin{aligned} \alpha\beta - \tfrac{e(S,T)}{|R|\Delta_R} &\geq \alpha\beta - (t-1)\beta/\Delta_R, \\ &= \beta(\alpha - (t-1)/\Delta_R). \end{aligned}$$

Lemma 58 implies that $\alpha\beta - e(S,T)/|R|\Delta_R \leq \mu\sqrt{\alpha\beta(1-\alpha)(1-\beta)}$. Since we expect α to be small, we will drop the $(1-\alpha)$ term, and we have

$$\begin{aligned} \mu\sqrt{\alpha\beta(1-\beta)} &\geq \beta(\alpha - (t-1)/\Delta_R), \\ \mu^2\alpha\beta(1-\beta) &\geq \beta^2(\alpha - (t-1)/\Delta_R)^2, \\ \mu^2(1-\beta)/\beta &\geq \alpha - 2(t-1)/\Delta_R + (t-1)^2/(\Delta_R^2\alpha). \end{aligned}$$

By hypothesis, $\alpha \geq (1+\epsilon)(t-1)/\Delta_R$. Let $c \geq 1$ such that $\alpha = c(1+\epsilon)(t-1)/\Delta_R$. Then,

$$\begin{aligned} \mu^2(1-\beta)/\beta &\geq c(1+\epsilon)(t-1)/\Delta_R - 2(t-1)/\Delta_R + (t-1)/(\Delta_R(1+\epsilon)c) \\ &= \left(c(1+\epsilon) - 2 + \tfrac{1}{c(1+\epsilon)}\right)\tfrac{t-1}{\Delta_R}. \end{aligned}$$

Define $f(x) = x(1+\epsilon) + x^{-1}(1+\epsilon)^{-1} - 2$ for $x \geq 1$. The derivative of $f(x)$ is

$$f'(x) = 1 + \epsilon - (1+\epsilon)^{-1}x^{-2}.$$

Since $1 + \epsilon > 1$, for any $x \geq 1$, we have $f'(x) > 0$. Hence, $f(c) \geq f(1)$, and

$$\begin{array}{rcl}
\mu^2(1-\beta)/\beta & \geq & (1+\epsilon-2+(1+\epsilon)^{-1})\frac{t-1}{\Delta_R}, \\
 & = & \frac{((\epsilon-1)(1+\epsilon)+1)(t-1)}{(1+\epsilon)\Delta_R}, \\
 & = & \frac{\epsilon^2(t-1)}{(1+\epsilon)\Delta_R}, \\
1/\beta - 1 & \geq & \frac{\epsilon^2(t-1)}{(1+\epsilon)\mu^2\Delta_R}, \\
\beta & \leq & \frac{(1+\epsilon)\mu^2\Delta_R}{\epsilon^2(t-1)+(1+\epsilon)\mu^2\Delta_R}, \\
1-\beta & \geq & \frac{\epsilon^2(t-1)}{\epsilon^2(t-1)+(1+\epsilon)\mu^2\Delta_R}.
\end{array}$$

Recall that $1 - \beta = |\Gamma_t(S)|/|R|$, so this completes the proof. □

4.4 Proof of Incidence Bounds

In this section, we prove Theorems 51, 53, and 54. We first prove results on the spectrum of the bipartite graph associated to a BIBD. We then use Lemmas 58 and 59 to complete the proofs.

Lemma 60. *Let (X, B) be an (r, k, λ)-BIBD. Let $G = (X, B, E)$ be a bipartite graph with left vertices X, right vertices B, and $(x, b) \in E$ if $x \in b$. Let A be the $(|X| + |B|) \times (|X| + |B|)$ adjacency matrix of G. Then, the normalized second eigenvalue of A is $\sqrt{(r-\lambda)/rk}$.*

Proof. Let N be the $|X| \times |B|$ incidence matrix of $\mathcal{D}$; that is, N is a $(0,1)$-valued matrix such that $N_{i,j} = 1$ iff point i is in block j. We can write

$$A = \begin{bmatrix} 0 & N \\ N^T & 0 \end{bmatrix}.$$

Instead of analyzing the eigenvalues of A directly, we'll first consider the eigenvalues of A^2. Since

$$A^2 = \begin{bmatrix} NN^T & 0 \\ 0 & N^T N \end{bmatrix}$$

is a block diagonal matrix, the eigenvalues of A^2 (counted with multiplicity) are the union of the eigenvalues of NN^T and the eigenvalues of $N^T N$. We will start by calculating the eigenvalues of NN^T.

The following observation about NN^T was noted by Bose [10].

Proposition 61.

$$NN^T = (r - \lambda)I + \lambda J,$$

where I is the $|X| \times |X|$ identity matrix and J is the $|X| \times |X|$ all-1s matrix.

Proof. The entry $(NN^T)_{i,j}$ corresponds to the number of blocks that contain both point i and point j. From the definition of an (r, k, λ)-BIBD, it follows that $(NN^T)_{i,i} = r$ and $(NN^T)_{i,j} = \lambda$ if $i \neq j$, and the conclusion of the proposition follows. $\square$

We use the above decomposition to calculate the eigenvalues of NN^T.

Proposition 62. *The eigenvalues of NN^T are rk with multiplicity 1 and $r - \lambda$ with multiplicity $|X| - 1$.*

Proof. The eigenvalues of I are all 1. The eigenvalues of J are $|X|$ with multiplicity 1 and 0 with multiplicity $|X| - 1$. The eigenvector of J corresponding to eigenvalue $|X|$ is the all-ones vector, and the orthogonal eigenspace has eigenvalue 0. Since I and J share a basis of eigenvectors, the eigenvalues of NN^T are simply the sums of the corresponding eigenvalues of $(r - \lambda)I$ and λJ. Hence, the largest eigenvalue of NN^T is $r - \lambda + |X|\lambda$, corresponding to the all-ones vector, and the remaining eigenvalues are $r - \lambda$, corresponding to vectors whose entries sum to 0. From equation (2), we have $\lambda(|X| - 1) = r(k - 1)$, and so we can write the largest eigenvalue as rk. $\square$

Next, we use the existence of a singular value decomposition of N to show that the nonzero eigenvalues of $N^T N$ have the same values and occur with the same multiplicity as the eigenvalues of NN^T.

The following is a standard theorem from linear algebra; see e.g. [53, p. 429].

Theorem 63 (Singular value decomposition.). *Let M be a $m \times n$ real-valued matrix with rank r. Then,*

$$M = P\Sigma Q^T,$$

where P is an $m \times m$ orthogonal matrix, Q is an $n \times n$ orthogonal matrix, and Σ is a diagonal matrix. In addition, if the diagonal entries of Σ are $s_1, s_2, \ldots, s_r, 0, \ldots, 0$, then the nonzero eigenvalues of MM^T and M^TM are $s_1^2, s_2^2, \ldots, s_r^2$.

It is immediate from this theorem that the nonzero eigenvalues of N^TN, counted with multiplicity, are identical with those of NN^T. Hence, the nonzero eigenvalues of A^2 are rk with multiplicity 2 and $r - \lambda$ with multiplicity $2(|X| - 1)$.

Clearly, the eigenvalues of A^2 are the squares of the eigenvalues of A; indeed, if x is an eigenvector of A with eigenvalue μ, then $A^2x = \mu Ax = \mu^2 x$. Hence, the conclusion of the lemma will follow from the following proposition that the eigenvalues of A are symmetric about 0. Although it is a well-known fact that the eigenvalues of the adjacency matrix of a bipartite graph are symmetric about 0, we include a simple proof here for completeness.

Proposition 64. *If μ is an eigenvalue of A with multiplicity w, then $-\mu$ is an eigenvalue of A with multiplicity w.*

Proof. Let $x_1 \in \mathbb{R}^{|X|}$ and $x_2 \in \mathbb{R}^{|B|}$ so that $(x_1, x_2)^T$ is an eigenvector of A with corresponding nonzero eigenvalue μ.

Then,

$$A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} Nx_2 \\ N^T x_1 \end{bmatrix} = \mu \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Note that, since $Nx_2 = \mu x_1$ and $N^T x_1 = \mu x_2$ and $\mu \neq 0$, we have that that $x_1 \neq 0$ and $x_2 \neq 0$.

In addition,

$$A \begin{bmatrix} -x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} Nx_2 \\ -N^T x_1 \end{bmatrix} = -\mu \begin{bmatrix} -x_1 \\ x_2 \end{bmatrix}$$

Hence, if μ is an eigenvalue of A with eigenvector $(x_1, x_2)^T$, then $-\mu$ is an eigenvalue of A with eigenvector $(-x_1, x_2)^T$. Since A is a real symmetric matrix, it follows from the spectral theorem (e.g. [53, p. 227]) that A has an orthogonal eigenvector basis; hence, we can match the eigenvectors of A with eigenvalue μ with those having eigenvalue $-\mu$ to show that the multiplicity of μ is equal to the multiplicity of $-\mu$. $\square$

Now we can calculate that the nonzero eigenvalues of A are $\sqrt{rk}$ and $-\sqrt{rk}$, each with multiplicity 1, and $\sqrt{r-\lambda}$ and $-\sqrt{r-\lambda}$, each with multiplicity $|X|-1$. Hence, the normalized second eigenvalue of A is $\sqrt{(r-\lambda)/rk}$, and the proof of Lemma 60 is complete. $\square$

Proof of Theorem 51. Lemma 58 implies that given a sets $P \subseteq X$ and $L \subseteq B$, the number of edges in G between P and L is bounded by

$$\left| \frac{e(P,L)}{r|X|} - \frac{|P|\,|L|}{|X|\,|B|} \right| \leq \sqrt{(r-\lambda)|P|\,|L|/rk|X|\,|B|}.$$

From equation (1), we know that $r|X| = k|B|$, so multiplying through by $r|X|$ gives

$$\big| e(P,L) - |P|\,|L|r/|B| \big| \leq \sqrt{(r-\lambda)|P|\,|L|}.$$

From the construction of G, we see that $e(P,L)$ is exactly the term $I(P,L)$ bounded in Theorem 51, so this completes the proof of Theorem 51. $\square$

Proof of Theorem 53. Lemma 59 implies that, for any $\epsilon \in \mathbb{R}_{>0}$ and $t \in \mathbb{Z}_{\geq 1}$, given a set $P \subseteq X$ such that

$$|P| \geq (1+\epsilon)(t-1)|X|/k,$$

there are at least

$$\Gamma_k(P) \geq \frac{\epsilon^2(t-1)|B|}{\epsilon^2(t-1) + (r-\lambda)k(1+\epsilon)/rk}$$

vertices in B that each have at least t edges to vertices in P. Rearranging slightly and again using the fact that edges in G correspond to incidences in $\mathcal{D}$ gives Theorem 53. $\square$

Proof of Theorem 54. Lemma 59 also implies that, for any $\epsilon \in \mathbb{R}_{>0}$ and $t \in \mathbb{Z}_{\geq 1}$, given a set $L \subseteq B$ such that

$$|L| \geq (1+\epsilon)(t-1)|B|/r = (1+\epsilon)(t-1)|X|/k,$$

we have

$$\Gamma_k(L) \geq \frac{\epsilon^2(t-1)|X|}{\epsilon^2(t-1)+(r-\lambda)r(1+\epsilon)/rk}.$$

Simplifying this expression gives Theorem 54. $\square$

4.5 Application to Distinct Triangle Areas

In this section, we will prove Theorem 57. We will need the following theorem, which was proved by Iosevich, Rudnev, and Zhai [39] as a key ingredient of their lower bound on distinct triangle areas.

Theorem 65 ([39]). *Let $F, G \subset \mathbb{F}_q^2$. Suppose $0 \notin F$. Let, for $d \in \mathbb{F}_q$,*

$$\nu(d) = |\{(a,b) \in F \times G : a \cdot b = d\}|,$$

where $a \cdot b = a_x b_x + a_y b_y$. Then

$$\sum_d \nu^2(d) \leq |F|^2|G|^2q^{-1} + q|F||G| \max_{x \in \mathbb{F}_q^2 \setminus \{0\}} |F \cap l_x|,$$

where

$$l_x = \{sx : s \in \mathbb{F}_q\}.$$

We will also need the following consequence of Corollary 55.

Lemma 66. *Let $\epsilon \in \mathbb{R}_{>0}$ and $t \in \mathbb{Z}_{\geq 2}$. There exists a constant $c'_\epsilon > 0$, depending only on ϵ, such that the following holds.*

Let P be a set of $(1+\epsilon)(t-1)q$ points in $\mathbb{F}_q^2$. Then there is a point $z \in P$ such that $c'_\epsilon q$ or more t-rich lines are incident to z. Moreover, if $\epsilon \geq 1$, then we can take $c'_\epsilon = 1/3$.

Proof. By Corollary 55,

$$|\Gamma_t(P)| \geq a_{\epsilon,t} q^2,$$

where

$$a_{\epsilon,t} = \frac{\epsilon^2(t-1)}{\epsilon^2(t-1)+1+\epsilon}.$$

Denote by $I(P, \Gamma_t(P))$ the number of incidences between points of P and lines of $\Gamma_t(P)$. Since each line of $\Gamma_t(P)$ is incident to at least t points of P, the average number of incidences with lines of $\Gamma_t(P)$ that each point of P participates in is at least

$$\begin{aligned} I(P, \Gamma_t(P))/|P| &\geq t|\Gamma_t(P)|/|P|, \\ &\geq t a_{\epsilon,t} q^2/|P|, \\ &= t a_{\epsilon,t} q/((1+\epsilon)(t-1)), \\ &= c'_{\epsilon,t} q, \end{aligned}$$

where

$$\begin{aligned} c'_{\epsilon,t} &= t a_{\epsilon,t}/((1+\epsilon)(t-1)), \\ &= t\epsilon^2/((1+\epsilon)(\epsilon^2(t-1)+1+\epsilon)). \end{aligned}$$

The derivative of $c'_{\epsilon,t}$ with respect to t is

$$\frac{\delta c'_{\epsilon,t}}{\delta t} = \frac{\epsilon^2(-\epsilon^2+\epsilon+1)}{(\epsilon+1)((t-1)\epsilon^2+\epsilon+1)^2}.$$

Since this derivative is positive for $0 < \epsilon < (1+\sqrt{5})/2$ and $t > 0$, we have that $c'_{\epsilon,t}$ is a monotonically increasing function of t for any fixed $0 < \epsilon < (1+\sqrt{5})/2$. Hence, for $\epsilon \leq 1$,

$$I(P, \Gamma_t(P))/|P| \geq c'_{\epsilon,2} q.$$

For $\epsilon \leq 1$, let $c'_\epsilon = c'_{\epsilon,2} = 2\epsilon^2/((1+\epsilon)(\epsilon^2+\epsilon+1))$. Since the expected number of t-rich lines incident to a point $p \in P$ chosen uniformly at random is at least $c'_\epsilon q$, there must be a point incident to so many t-rich lines.

If $\epsilon > 1$, choose an arbitrary set $P' \subset P$ of size $|P'| = 2(t-1)q$. By the preceding argument, there must be a point $z \in P'$ incident to at least $c'_1 q = 1/3q$ lines that are t-rich in P'. Hence, z is also incident to at least so many lines that are t-rich in P.

$\square$

Proof of Theorem 57. Let $\delta = (1+\epsilon)/(t-1) - 1$, so that $|P| = (1+\delta)(t-1)q$. Let c'_δ be as in Lemma 66. By Lemma 66, there is a point z in P incident to $c'_\delta q$ or more t-rich lines. Let $P' \subseteq P - \{z\}$ be a set of points such that there are exactly $t-1$ points of P' on exactly $\lceil c'_\delta q \rceil$ lines incident to z. Clearly,

$$|P'| = (t-1)\lceil c'_\delta q \rceil \geq (t-1)c'_\delta q.$$

Let P'_z be P' translated so that z is at the origin.

Each ordered pair $(a,b) \in P'_z \times P'_z$ corresponds to a triangle having z as a vertex. By the definition of area, given in Section 4.2.3, the area of the triangle corresponding to (a,b) is $a_x b_y - b_x a_y$. For any point $x \in \mathbb{F}_q^2$, let $x^\perp = (-x_y, x_x)$; let $P'^\perp_z = \{x^\perp : x \in P'_z\}$. The area corresponding to (a,b) is $a_x b_y - b_x a_y = a^\perp \cdot b$.

Hence, the number of distinct areas spanned by triangles with z as a vertex is at least the number of distinct dot products $|\{a^\perp \cdot b : a^\perp \in P'^\perp_z, b \in P'_z\}|$. To write this in another way, let $\nu(d)$ be as defined in Theorem 65 with $F = P'_z$ and $G = P'^\perp_z$. Then, the number of distinct areas spanned by triangles containing z is at least $|\{d : \nu(d) \neq 0\}|$.

Since no line through the origin contains more than $t-1$ points of P'_z, Theorem 65 implies that

$$\sum_d \nu^2(d) \leq |P'|^4 q^{-1} + q(t-1)|P'|^2.$$

By Cauchy-Schwarz, the number of distinct triangle areas is at least

$$\begin{aligned}
|\{d : \nu(d) \neq 0\}| &\geq |\textstyle\sum_d \nu(d)|^2 \left(\sum_d \nu^2(d)\right)^{-1} \\
&= |\{(x,y) \in F \times G\}|^2 \left(\textstyle\sum_d \nu^2(d)\right)^{-1} \\
&\geq |P'|^4 \left(|P'|^4 q^{-1} + q|P'|^2(t-1)\right)^{-1} \\
&= q\left(1 + q^2(t-1)|P'|^{-2}\right)^{-1} \\
&\geq q\left(1 + (t-1)^{-1} c'^{-2}_\delta\right)^{-1} \\
&= \left((t-1)c'^2_\delta/((t-1)c'^2_\delta + 1)\right) q.
\end{aligned}$$

Hence, P includes a point z that is a vertex of triangles with at least

$$c_\epsilon q = \max_t \left(((t-1){c'_\delta}^2)/((t-1){c'}^2_\delta + 1) \right) q$$

distinct areas. To complete the proof, check that c_ϵ has the claimed properties that $c_\epsilon > 0$ for any ϵ, and that $c_\epsilon \to 1$ as $\epsilon \to \infty$. □

4.6 Proof of Lemma 58

The proof here follows closely the proof of Lemma 4.15 in [63].

Proof. Let χ_S be the characteristic row vector of S in L; in other words, χ_S is a vector of length $|L|$ with entries in $\{0,1\}$ such that $(\chi_S)_i = 1$ iff vertex i is in S. Similarly, let χ_T be the characteristic vector of T in R. Note that

$$e(S,T) = \chi_S A \chi_T^t, \tag{4.3}$$

where χ_T^t is the transpose of χ_T.

Let $U_L = (|L|^{-1}, |L|^{-1}, \ldots, |L|^{-1})$ be the uniform distribution on L, and let $U_R = (|R|^{-1}, |R|^{-1}, \ldots, |R|^{-1})$ be the uniform distribution on R. We can express χ_S as the sum of a component parallel to U_L and $\chi_S^\perp$ orthogonal to U_L.

$$\begin{aligned} \chi_S &= (\langle \chi_S, U_L \rangle / \langle U_L, U_L \rangle)\, U_L + \chi_S^\perp \\ &= \textstyle\sum_i (\chi_S)_i U_L + \chi_S^\perp \\ &= \alpha |L| U_L + \chi_S^\perp. \end{aligned}$$

Similarly, let $\chi_T^\perp$ be a vector orthogonal to U_R so that

$$\chi_T = \beta |R| U_R + \chi_T^\perp.$$

From equation 4.3, we have

$$\begin{aligned} e(S,T) &= \chi_S A \chi_T^t, \\ &= \left(\alpha|L|U_L + \chi_S^\perp\right) A \left(\beta|R|U_R + \chi_T^\perp\right)^t, \\ &= \alpha\beta|L||R|U_L A U_R^t + \chi_S^\perp A U_R^t + U_L A(\chi_T^\perp)^t + \chi_S^\perp A(\chi_T^\perp)^t. \end{aligned}$$

From the definitions, we can calculate that

$$\begin{aligned} U_L A &= |L|^{-1}\Delta_R|R|U_R &= \Delta_L U_R, \\ A U_R^t &= |R|^{-1}\Delta_L|L|U_L^t &= \Delta_R U_L^t, \\ U_L U_L^t &= |L|^{-1}, \\ U_R U_R^t &= |R|^{-1}. \end{aligned}$$

Combined with the orthogonality of $\chi_S^\perp$ with U_L and of $\chi_T^\perp$ with U_R, we have

$$\begin{aligned} e(S,T) &= \alpha\beta|L|\Delta_L + \chi_S^\perp A(\chi_T^\perp)^t, \\ &= \alpha\beta \cdot e(G) + \chi_S^\perp A(\chi_T^\perp)^t. \end{aligned}$$

Hence,

$$\begin{aligned} \left|\tfrac{e(S,T)}{e(G)} - \alpha\beta\right| &= \left|(\chi_S^\perp A)(\chi_T^\perp)^t/(|L|\Delta_L)\right| \\ &\leq \|\chi_S^\perp A\|\|\chi_T^\perp\|/(|L|\Delta_L) \\ &\leq \mu_2\|\chi_S^\perp\|\|\chi_T^\perp\|/(|L|\Delta_L) \end{aligned}$$

The trivial eigenvalue of a (Δ_L, Δ_R) biregular, bipartite graph is $\sqrt{\Delta_L\Delta_R}$; hence, $\mu = \mu_2/\sqrt{\Delta_L, \Delta_R}$, and so

$$\left|\frac{e(S,T)}{e(G)} - \alpha\beta\right| \leq \mu\sqrt{\frac{\Delta_R}{|L|^2\Delta_L}}\|\chi_S^\perp\|\|\chi_T^\perp\| = \mu\sqrt{\frac{1}{|L||R|}}\|\chi_S^\perp\|\|\chi_T^\perp\|. \qquad (4.4)$$

Note that

$$\alpha|L| = \|\chi_S\|^2 = \|\alpha|L|U_L\|^2 + \|\chi_S^\perp\|^2 = \alpha^2|L| + \|\chi_S^\perp\|^2,$$

so

$$\|\chi_S^\perp\| = \sqrt{\alpha(1-\alpha)|L|}.$$

Similarly,

$$\|\chi_T^{\perp}\| = \sqrt{\beta(1-\beta)|R|}.$$

Substituting these equalities into expression (4.4) completes the proof. $\square$

References

[1] Oswin Aichholzer and Franz Aurenhammer. Classifying hyperplanes in hypercubes. *SIAM Journal on Discrete Mathematics*, 9(2):225–232, 1996.

[2] Gerald L Alexanderson and John E Wetzel. A simplicial 3-arrangement of 21 planes. *Discrete mathematics*, 60:67–73, 1986.

[3] Noga Alon. Eigenvalues, geometric expanders, sorting in rounds, and Ramsey theory. *Combinatorica*, 6(3):207–219, 1986.

[4] Noga Alon and Fan RK Chung. Explicit construction of linear sized tolerant networks. *Discrete Mathematics*, 72(1):15–19, 1988.

[5] Boris Aronov and Micha Sharir. Cutting circles into pseudo-segments and improved bounds for incidences. *Discrete & Computational Geometry*, 28(4):475–490, 2002.

[6] Simeon Ball and Zsuzsa Weiner. An introduction to finite geometry. http://www-ma4.upc.es/ simeon/IFG.pdf.

[7] József Beck. On the lattice property of the plane and some problems of Dirac, Motzkin and Erdős in combinatorial geometry. *Combinatorica*, 3(3-4):281–297, 1983.

[8] Mike Bennett, Alex Iosevich, and Jonathan Pakianathan. Three-point configurations determined by subsets of $\mathbb{F}_q^2$ via the Elekes-Sharir paradigm. *Combinatorica*, 34(6):689–706, 2014.

[9] Bela Bollobas. *Graph theory: an introductory course*, volume 63. Springer Science & Business Media, 2012.

[10] RC Bose. A note on Fisher's inequality for balanced incomplete block designs. *The Annals of Mathematical Statistics*, 20(4):619–620, 1949.

[11] Jean Bourgain, Nets Katz, and Terence Tao. A sum-product estimate in finite fields, and applications. *Geometric & Functional Analysis GAFA*, 14(1):27–57, 2004.

[12] Peter Brass, William OJ Moser, and János Pach. *Research problems in discrete geometry*. Springer Science & Business Media, 2005.

[13] Mei-Chu Chang and József Solymosi. Sum-product theorems and incidence geometry. *J. Eur. Math. Soc. (JEMS)*, 9(3):545–560, 2007.

[14] Javier Cilleruelo. Combinatorial problems in finite fields and Sidon sets. *Combinatorica*, 32(5):497–511, 2012.

[15] David Covert, Derrick Hart, Alex Iosevich, Doowon Koh, and Misha Rudnev. Generalized incidence theorems, homogeneous forms and sum-product estimates in finite fields. *European J. Combin.*, 31(1):306–319, 2010.

[16] Hallard T Croft, Kenneth J Falconer, and Richard K Guy. *Unsolved Problems in Geometry: Unsolved Problems in Intuitive Mathematics*, volume 2. Springer Science & Business Media, 2012.

[17] N. G. de Bruijn and P. Erdős. On a combinatorial problem. *Nederl. Akad. Wetensch., Proc.*, 51:1277–1279 = Indagationes Math. 10, 421–423, 1948.

[18] Nicolaas G de Bruijn and Paul Erdős. On a combinatorial problem. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen Indagationes mathematicae*, 51(105):1277–1277, 1948.

[19] Thao Do. Extending Erdös-Beck's theorem to higher dimensions. *arXiv preprint arXiv:1607.00048*, 2016.

[20] Zeev Dvir. Incidence theorems and their applications. *Found. Trends Theor. Comput. Sci.*, 6(4):257–393 (2012), 2010.

[21] György Elekes. On the number of sums and products. *Acta Arith.*, 81(4):365–367, 1997.

[22] György Elekes and Micha Sharir. Incidences in three dimensions and distinct distances in the plane. *Combinatorics, Probability and Computing*, 20(04):571–608, 2011.

[23] György Elekes and Csaba D Tóth. Incidences of not-too-degenerate hyperplanes. In *Proceedings of the twenty-first annual symposium on Computational geometry*, pages 16–21. ACM, 2005.

[24] Jordan S Ellenberg and Marton Hablicsek. An incidence conjecture of Bourgain over fields of positive characteristic. *arXiv preprint arXiv:1311.1479*, 2013.

[25] Paul Erdős. On sets of distances of n points. *The American Mathematical Monthly*, 53(5):248–250, 1946.

[26] Pál Erdős. On the combinatorial problems which i would most like to see solved. *Combinatorica*, 1(1):25–42, 1981.

[27] Paul Erdős, G Purdy, and Ernst G Straus. On a problem in combinatorial geometry. *Discrete Mathematics*, 40(1):45–52, 1982.

[28] Paul Erdős and George Purdy. Extremal problems in combinatorial geometry. In *Handbook of combinatorics (vol. 1)*, pages 809–874. MIT Press, 1996.

[29] Jacob Fox, János Pach, Adam Sheffer, Andrew Suk, and Joshua Zahl. A semi-algebraic version of Zarankiewicz's problem. *arXiv preprint arXiv:1407.5705*, 2014.

[30] Rota Gian-Carlo. Combinatorial theory, old and new. In *Proceedings of the International Mathematical Congress Held...*, volume 3, page 229. University of Toronto Press, 1971.

[31] Codrut Grosu. $\mathbb{F}_p$ is locally like $\mathbb{C}$. *Journal of the London Mathematical Society*, 89(3):724–744, 2014.

[32] Branko Grünbaum. A catalogue of simplicial arrangements in the real projective plane. *Ars Mathematica Contemporanea*, 2(1), 2009.

[33] Branko Grünbaum and Geoffrey C Shephard. Simplicial arrangements in projective 3-space. *Mitt. Math. Semin. Giessen*, 166:49–101, 1984.

[34] Larry Guth and Nets Hawk Katz. Algebraic methods in discrete analogs of the Kakeya problem. *Advances in Mathematics*, 225(5):2828–2839, 2010.

[35] Larry Guth and Nets Hawk Katz. On the Erdős distinct distances problem in the plane. *Annals of Mathematics*, 181(1):155–190, 2015.

[36] Brandon Hanson, Ben Lund, and Oliver Roche-Newton. On distinct perpendicular bisectors and pinned distances in finite fields. *Finite Fields and Their Applications*, 37:240–264, 2016.

[37] Terry Tao (http://mathoverflow.net/users/766/terry tao). Where did the term "additive energy" originate? MathOverflow. URL:http://mathoverflow.net/q/223962 (version: 2015-11-18).

[38] Alex Iosevich, S Konyagin, Michael Rudnev, and V Ten. Combinatorial complexity of convex sequences. *Discrete & Computational Geometry*, 35(1):143–158, 2006.

[39] Alex Iosevich, Misha Rudnev, and Yujia Zhai. Areas of triangles and becks theorem in planes over finite fields. *Combinatorica*, pages 1–14, 2012.

[40] Timothy GF Jones. Further improvements to incidence and Beck-type bounds over prime finite fields. *arXiv preprint arXiv:1206.4517*, 2012.

[41] Nets Hawk Katz and Gábor Tardos. A new entropy inequality for the erdos distance problem. *Contemporary Mathematics*, 342:119–126, 2004.

[42] Izabella Łaba. From harmonic analysis to arithmetic combinatorics. *Bulletin (New Series) of the American Mathematical Society*, 45(1):77–115, 2008.

[43] Ben Lund. Essential dimension and the flats spanned by a point set. *arXiv preprint arXiv:1602.08002*, 2016.

[44] Ben Lund. A refined energy bound for perpendicular bisectors. *arXiv preprint arXiv:1604.02059*, 2016.

[45] Ben Lund and Shubhangi Saraf. Incidence bounds for block designs. *SIAM Journal on Discrete Mathematics*, 30(4):1997–2010, 2016.

[46] Ben Lund, Adam Sheffer, and Frank de Zeeuw. Bisector energy and few distinct distances. *Discrete & Computational Geometry*, pages 1–20.

[47] Ben D Lund, George B Purdy, and Justin W Smith. A bichromatic incidence bound and an application. *Discrete & Computational Geometry*, 46(4):611–625, 2011.

[48] John H Mason. Matroids: Unimodal conjectures and Motzkins theorem. *Combinatorics (D. JA Welsh and DR Woodall, eds.), Institute of Math. and Appl*, pages 207–221, 1972.

[49] János Pach and Micha Sharir. Geometric incidences. In *Towards a theory of geometric graphs*, volume 342 of *Contemp. Math.*, pages 185–223. Amer. Math. Soc., Providence, RI, 2004.

[50] János Pach and Micha Sharir. Combinatorial geometry with algorithmic applications. *The Alcala Lectures, Alcala, Spain*, 2006.

[51] Rom Pinchasi. The minimum number of distinct areas of triangles determined by a set of n points in the plane. *SIAM Journal on Discrete Mathematics*, 22(2):828–831, 2008.

[52] George Purdy. Two results about points, lines and planes. *Discrete mathematics*, 60:215–218, 1986.

[53] Steven Roman. *Advanced linear algebra, volume 135 of Graduate Texts in Mathematics*. Springer-Verlag, New York, 1992.

[54] H. J. Ryser. An extension of a theorem of de Bruijn and Erdős on combinatorial designs. *J. Algebra*, 10:246–261, 1968.

[55] József Solymosi and Terence Tao. An incidence theorem in higher dimensions. *Discrete & Computational Geometry*, 2(48):255–280, 2012.

[56] József Solymosi and Cs D Tóth. Distinct distances in the plane. *Discrete & Computational Geometry*, 25(4):629–634, 2001.

[57] Sophie Stevens and Frank de Zeeuw. An improved point-line incidence bound over arbitrary fields. *arXiv preprint arXiv:1609.06284*, 2016.

[58] E. Szemerédi and W. T. Trotter. Extremal problems in discrete geometry. *Combinatorica*, 3(3):381–392, 1982.

[59] Endre Szemerédi and William T Trotter Jr. Extremal problems in discrete geometry. *Combinatorica*, 3(3-4):381–392, 1983.

[60] Terence Tao. Expanding polynomials over finite fields of large characteristic, and a regularity lemma for definable sets. *arXiv preprint arXiv:1211.2894*, 2012.

[61] Csaba D Tóth. The Szemerédi-Trotter theorem in the complex plane. *arXiv preprint math/0305283*, 2003.

[62] Csaba D Tóth. The Szemerédi-Trotter theorem in the complex plane. *Combinatorica*, 35(1):95–126, 2015.

[63] Salil P Vadhan. *Pseudorandomness*. Now, 2012.

[64] Le Anh Vinh. The Szemerédi–Trotter type theorem and the sum-product estimate in finite fields. *European Journal of Combinatorics*, 32(8):1177–1181, 2011.

[65] Le Anh Vinh. On point-line incidences in vector spaces over finite fields. *Discrete applied mathematics*, 177:146–151, 2014.

[66] Hassler Whitney. Elementary structure of real algebraic varieties. *Annals of Mathematics*, pages 545–556, 1957.

[67] Joshua Zahl. A Szemerédi-Trotter type theorem in $\mathbb{R}^4$. *Discrete & Computational Geometry*, 54(3):513–572, 2015.