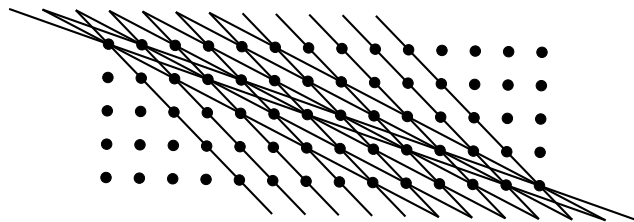# Polynomial Methods and Incidence Theory

Adam Sheffer

This document is an incomplete draft from July 26, 2020.
Several chapters are still missing.

Acknowledgements

Contents

Introduction

> *"Algebra is the offer made by the devil to the mathematician. The devil says: I will give you this powerful machine, it will answer any question you like. All you need to do is give me your soul: give up geometry and you will have this marvellous machine."* / Michael Atiyah [3].

In Paul Halmos's famous essay on how to write mathematics [53], Halmos writes "Just as there are two ways for a sequence not to have a limit (no cluster points or too many), there are two ways for a piece of writing not to have a subject (no ideas or too many)." The book you are now starting has two main subjects, which is hopefully a reasonable amount. These two subjects, *the polynomial method* and *incidence theory*, are closely tied and hard to separate.

Geometric incidences are a family of combinatorial problems, which existed for many decades as part of discrete geometry. In the past decade, incidence problems have been experiencing a renaissance. New interesting connections between incidences and other parts of mathematics are constantly being exposed (such as harmonic analysis, theoretical computer science, model theory, and number theory). At the same time, significant progress is being made on long-standing open incidence problems. The study of geometric incidences is currently an active and exciting research field. One purpose of this book is to survey this field, the recent developments in it, and a variety of connections to other fields.
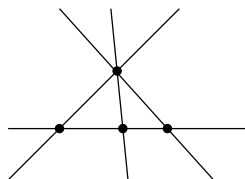
Figure 1: A configuration of four points, four lines, and nine incidences.

In an incidence problem we have a set of points $\mathcal{P}$ and a set of geometric objects $\mathcal{V}$. An incidence is a pair $(p, V) \in \mathcal{P} \times \mathcal{V}$ such that the point p is contained in the object

V. We denote by $I(\mathcal{P}, \mathcal{V})$ the number of incidences in $\mathcal{P} \times \mathcal{V}$, and (most commonly) wish to study the maximum value $I(\mathcal{P}, \mathcal{V})$ can have. One of the simplest incidence problems studies the maximum number of incidences between m points and n lines in the real plane (see Figure 1). Other variants include incidences with other types of curves, incidences with higher-dimensional algebraic objects in $\mathbb{R}^d$, and incidences with semi-algebraic sets in $\mathbb{R}^d$. Incidence problems are also being studied in $\mathbb{C}^d$, in spaces over finite fields, o-minimal structures, and more.

Much of the recent progress in studying incidence problems is due to new algebraic techniques. One may describe the philosophy behind these techniques as

> *Collections of objects that exhibit extremal behavior often have hidden algebraic structure. This algebraic structure can be exploited to gain a better understanding of the original problem.*

For example, in a point-line configuration that determines many incidences, one might expect the point set to have some sort of a lattice structure. Intuitively, one exposes the algebraic structure by defining polynomials according to the problem, and then studying properties of these polynomials. In an incidence problem, one might wish to study a polynomial that vanishes on the point set. This approach is often referred to as *the polynomial method*. In our study of incidences, we will focus on polynomial methods. In addition, we will see how polynomials methods are used to study problems that do not directly involve incidences.

The polynomial approach to studying incidence problems started around 2010. The field is still developing, and in some sense the foundations are not completely established yet. In particular, there are many interesting open problems, some which have not been thoroughly studied yet. Many chapters end by describing such open problems and conjectures. These are mostly long-standing difficult problems, and are meant to illustrate the current fronts of the field and the main difficulties we are currently facing.

Two other good sources for polynomial methods in Discrete Geometry are Guth's book "Polynomial Methods in Combinatorics" [49] and Dvir's survey "Incidence theorems and their applications" [27]. Although the current book and these two sources deal with similar topics, the overlap between them is smaller than one might expect and each has a different focus.

Glancing at the table of contents, one notices that some sections are defined as optional. Some optional sections contain standard technical parts of a proof which may not provide any new insights. Other optional sections require familiarity with a topic orthogonal to the main theme of the book. For example, the optional Sections 7.5 and 9.3 require basic familiarity with Differential Topology. The reason for

marking each section optional is explained in the relevant chapter.

Many results that are stated as *claims*, rather than as theorems or lemmas. These are results that seem worth stating but are too minor to call a theorem.

How to read this book

There are many ways to read this book, depending on the goal of the reader. One can of course read from cover to cover, but here are some other options.

- **A quick glance at Discrete Geometry.** For a brief introduction to incidences, other related Discrete Geometry problems, and basic classical proof techniques: read Chapter 1. This chapter does not involve any polynomial methods.

- **A basic polynomial incidence proof.** To understand how to derive incidence results using polynomial methods: read Chapters 1–3. Chapter 2 contains a minimal introduction to Algebraic Geometry in the real plane. Chapter 3 derives the basics of the *polynomial partitioning* technique, and uses this technique to prove basic incidence bounds in the real plane.

- **A variety of polynomial methods in combinatorics.** To see a variety of polynomial methods in combinatorics: read Chapters 1–6. In addition to proving incidence results using polynomial partitioning, Chapters 5 and 6 contain three other polynomial breakthroughs. Chapter 4 introduces more basic concepts from real Algebraic Geometry, this time in $\mathbb{R}^d$. Chapter 5 contains the joints theorem of Guth and Katz [50]. It is also a warmup for working in higher-dimensions without using polynomial partitioning. Chapter 6 contains two combinatorial results in finite fields: The finite field Kakeya theorem and the cap set problem.

- **The distinct distances theorem.** To understand the distinct distances theorem of Guth and Katz [51]: read Chapters 1–5 and 7–10, with Chapters 5, and 10 optional. Chapter 7 focuses on incidences in the complex plane. While doing that, the chapter also introduces the *constant-degree polynomial-partitioning* technique and uses it to derive incidence bounds in $\mathbb{R}^d$. Chapters 8 and 9 prove the distinct distances theorem. These chapters reduce the distances problem to an incidence problem in $\mathbb{R}^3$ and solve this incidence problem. Chapter 10 studies a couple of variants of the distinct distances problem.

- **Incidences in $\mathbb{R}^d$.** To understand advanced incidence techniques in $\mathbb{R}^d$: read Chapters 1–5 and Chapters 11–12. Chapter 11 contains more advanced techniques for deriving incidence bounds with real varieties of any dimension. Chapter 12 describes a few applications for incidence problems in $\mathbb{R}^d$.

- **Incidences in spaces over finite fields.** To study the most recent incidence results in finite fields: Read Chapters 6 and 13. It might be nice to get some context by first reading Chapter 1 or Chapters 1–3, but not necessary. The techniques in these chapters are very different than the ones in the rest of the book, and only require glancing at a couple of lemmas from previous chapters.

Figure 2 illustrates the chapter dependencies, and shows the various ways to read the book.
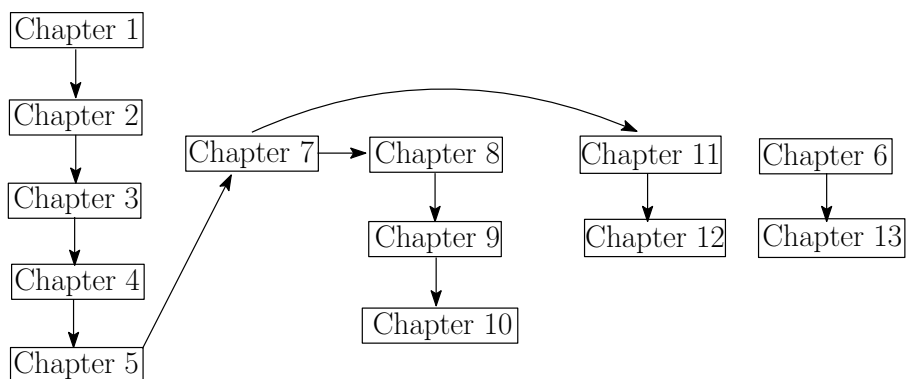
Figure 2: Chapter dependencies.

This book draft is not yet complete. A few chapters containing the most advanced topics are still missing.

Notation and inequalities

We use standard asymptotic notation. That is, $f(n) = O(g(n))$ implies that there exist constants c, n_0, such that for any $n \geq n_0$, we have $f(n) \leq c \cdot g(n)$. For example, $10n^2 + 1000 = O(n^2)$ holds since we can take $c = 100$ and $n_0 = 20$. Similarly, $f(n) = \Omega(g(n))$ implies that there exist constants c, n_0, such that for any $n \geq n_0$, we have $f(n) \geq c \cdot g(n)$. In addition, $f(n) = \Theta(g(n))$ implies that both $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$ hold. The notation $f(n) = o(g(n))$ implies that $f(n) = O(g(n))$ and $f(n) \neq \Theta(g(n))$ (that is, that $f(n)$ is asymptotically smaller than $g(n)$). When

writing an expression of the form $O_{s,t}(\cdot)$, we mean that the hidden constant may depend on the variables s and t. For example, $10s^{100}n^2 + s^{t^{100s}} = O_{s,t}(n^2)$.

We use standard graph theoretic notation. We usually denote a graph as $G = (V, E)$. We denote a bipartite graph as $G = (V \cup U, E)$, where V and U are the two vertex sets. For positive integers s and t, we define by $K_{s,t}$ the complete bipartite graph with s vertices on one side, t vertices on the other side, and all of the st edges between the two sides.

We denote the expectation of a random variable X as $\mathbb{E}[X]$. This is to prevent confusion between expectation and sets that are denoted as E.

Inequalities. We will use the *Cauchy-Schwarz inequality* rather often.

Theorem (The Cauchy-Schwarz Inequality). *Consider a positive integer n and two sequences of real numbers $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$. Then*

$$\sum_{i=1}^{n} |a_i b_i| \leq \sqrt{\left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right)}.$$

We will also rely on *Hölder's inequality*, which generalizes the Cauchy-Schwarz inequality.

Theorem (Hölder's Inequality). *Consider a positive integer n and two sequences of real numbers $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$. Let $1 < p, q$ satisfy $1/p + 1/q = 1$. Then*

$$\sum_{i=1}^{n} |a_i b_i| \leq \left(\sum_{i=1}^{n} a_i^p\right)^{1/p}\left(\sum_{i=1}^{n} b_i^q\right)^{1/q}.$$

Chapter 1

Incidences in Classical Discrete Geometry

> *"My most striking contribution to geometry is, no doubt, my problem on the number of distinct distances. This can be found in many of my papers on combinatorial and geometric problems."* / Paul Erdős, in a survey of his favorite contributions to mathematics, compiled for the celebrations of his 80'th birthday [38].

1.1   Introduction

In this chapter we introduce the concept of incidences, together with some first bounds and related problems. At this point we only discuss classical discrete geometry, from before the introduction of the polynomial method. This makes the current chapter rather different than the rest of the book (readers who prefer to avoid graph theory may wish to skip Sections 1.3–1.5). Nevertheless, some tricks that are introduced in this chapter are used throughout the book.

Given a set $\mathcal{P}$ of points and a set $\mathcal{L}$ of lines, both in $\mathbb{R}^2$, an *incidence* is a pair $(p, \ell) \in \mathcal{P} \times \mathcal{L}$ such that the point p is contained in the line ℓ. We denote by $I(\mathcal{P}, \mathcal{L})$ the number of incidences in $\mathcal{P} \times \mathcal{L}$. For example, Figure 1 (in the introduction of this book) depicts a configuration with nine incidences. For any m and n, Erdős constructed a set $\mathcal{P}$ of m points and a set $\mathcal{L}$ of n lines with $\Theta(m^{2/3}n^{2/3} + m + n)$ incidences. Erdős and Purdy [36] conjectured that no point-line configuration has an asymptotically larger number of incidences. This conjecture has been proven by Szemerédi and Trotter [97] in 1983.

Theorem 1.1 (The Szemerédi-Trotter theorem). *Let $\mathcal{P}$ be a set of m points and let $\mathcal{L}$ be a set of n lines, both in $\mathbb{R}^2$. Then $I(\mathcal{P}, \mathcal{L}) = O(m^{2/3}n^{2/3} + m + n)$.*

Szemerédi and Trotter's original proof is rather involved. In this chapter we present a later elegant proof by Székely [96]. A more general algebraic proof is presented in Chapter 3.

Finding the maximum number of incidences between points and lines in $\mathbb{R}^2$ is one of the simplest incidence problems, and almost the only one that is completely settled. Other problems involve incidences with circles or other types of curves, incidences with varieties in $\mathbb{R}^d$, with semi-algebraic objects in $\mathbb{R}^d$, in complex spaces $\mathbb{C}^d$, in spaces over finite fields, and many more. In each of these problems we wish to find the maximum number of incidences between a set of points and a set of geometric objects. This introductory chapter consists mostly of incidences with lines in $\mathbb{R}^2$.

One reason for studying incidence problems is that they are natural and elementary combinatorial problems. In this chapter we start to observe two additional reasons for studying incidence problems:

- *Incidence problems do not involve only combinatorial work, but also the study of the underlying geometry.* One example of this appears in Section 1.5, where we introduce the unit distances problem. As we will see, this problem involves studying properties that distinguish the Euclidean norm from almost all other distance norms.
- *Incidence results are useful also for problems that may not seem related to geometry.* In Section 1.8 we will see the sum-product problem, which started as a number theoretic problem not involving any geometry.

1.2   First proofs

The purpose of this section is to develop some initial intuition about incidences. We begin by deriving our first bound for an incidence problem. This is a weak bound and easy to prove. However, it is still useful in some cases (for example, see the proof of Lemma 1.15 below).

Lemma 1.2. *Let $\mathcal{P}$ be a set of m points and let $\mathcal{L}$ be a set of n lines, both in $\mathbb{R}^2$. Then $I(\mathcal{P}, \mathcal{L}) = O(m\sqrt{n} + n)$ and $I(\mathcal{P}, \mathcal{L}) = O(n\sqrt{m} + m)$.*

Proof. We only derive $I(\mathcal{P}, \mathcal{L}) = O(m\sqrt{n} + n)$. The other bound is obtained in a symmetric manner. If $I(\mathcal{P}, \mathcal{L}) < 2n$ then we are done, so we may assume that $I(\mathcal{P}, \mathcal{L}) \geq 2n$. Consider the set of triples

$$T = \{(a, b, \ell) \in \mathcal{P}^2 \times \mathcal{L} : a \text{ and } b \text{ are both incident to } \ell\}.$$

Let m_j be the number of points of $\mathcal{P}$ that are incident to the j'th line of $\mathcal{L}$. Note that $I(\mathcal{P}, \mathcal{L}) = \sum_{j=1}^{n} m_j$. Moreover, the number of triples of T that include the j'th line of $\mathcal{L}$ is exactly $\binom{m_j}{2}$. That is, $|T| = \sum_{j=1}^{n} \binom{m_j}{2}$. By applying the Cauchy-Schwarz inequality and recalling the assumption $I(\mathcal{P}, \mathcal{L}) \geq 2n$, we have

$$|T| = \sum_{j=1}^{n} \binom{m_j}{2} = \sum_{j=1}^{n} \frac{m_j(m_j - 1)}{2} = \frac{\sum_{j=1}^{n} m_j^2}{2} - \frac{\sum_{j=1}^{n} m_j}{2}$$

$$\geq \frac{I(\mathcal{P}, \mathcal{L})^2}{2n} - \frac{I(\mathcal{P}, \mathcal{L})}{2} = \Omega\left(\frac{I(\mathcal{P}, \mathcal{L})^2}{n}\right).$$

Since for every $a, b \in \mathcal{P}$ at most one line of $\mathcal{L}$ is incident to both a and b, we have $|T| = O(m^2)$. By combining the two bounds for $|T|$, we get that $I(\mathcal{P}, \mathcal{L})^2/n = O(m^2)$. Tidying this up yields $I(\mathcal{P}, \mathcal{L}) = O(m\sqrt{n})$. $\qquad\square$

In the proof of Lemma 1.2 we used a common combinatorial method called *double counting*. In this method we bound some quantity in two different ways, and then compare the two bounds to obtain new information about a different quantity. For example, in the above proof we counted the size of T in two different ways, and by comparing these two bounds we obtained a bound for the number of incidences. We will encounter this technique rather frequently in this book.

Note that in the proof of Lemma 1.2 we did not use any geometry beyond the observation that two lines intersect at most once. When replacing the lines with arbitrary sets that satisfy this property, the bound of the lemma becomes tight. Thus, to obtain the stronger Szemerédi-Trotter bound (Theorem 1.1) we need to rely on additional geometric properties of lines.

We now consider an asymptotically tight lower bound for Theorem 1.1. Instead of Erdős' original construction, we present a simpler construction due to Elekes [32].
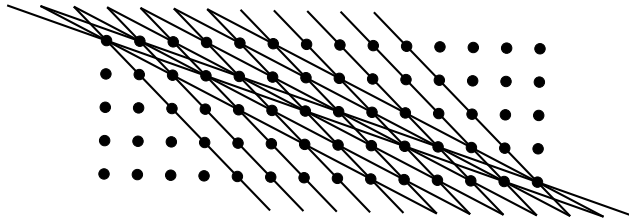
Figure 1.1: Elekes' construction, rotated by 90°.

Claim 1.3. *For every m and n there exist a set $\mathcal{P}$ of m points and a set $\mathcal{L}$ of n lines, both in $\mathbb{R}^2$, such that $I(\mathcal{P}, \mathcal{L}) = \Theta(m^{2/3}n^{2/3} + m + n)$.*

Proof. It is not difficult to check that the m term dominates the bound when $m = \Omega(n^2)$. In this case we can simply take m points on a single line, to obtain m incidences. Similarly, the n term dominates the bound when $n = \Omega(m^2)$. In this case we take n lines through a single point, to obtain n incidences. It remains to construct a configuration with $\Theta(m^{2/3}n^{2/3})$ incidences when $m = O(n^2)$ and $n = O(m^2)$.

Let $r = (m^2/4n)^{1/3}$ and $s = (2n^2/m)^{1/3}$ (for simplicity, instead of taking the ceiling function of s and r, we assume that these are integers). We set

$$\mathcal{P} = \{ (i,j) \ : \ 1 \le i \le r \quad \text{and} \quad 1 \le j \le 2rs \},$$

and

$$\mathcal{L} = \{ y = ax + b \ : \ 1 \le a \le s \quad \text{and} \quad 1 \le b \le rs \}.$$

Note that this construction consists of a rectangular section of the integer lattice and of a "lattice" of lines; such a configuration is depicted in Figure 1.1, rotated by $90°$. Also, notice that we indeed have

$$|\mathcal{P}| = 2r^2 s = 2 \cdot \frac{m^{4/3}}{(4n)^{2/3}} \cdot \frac{(2n^2)^{1/3}}{m^{1/3}} = m,$$

and

$$|\mathcal{L}| = rs^2 = \frac{m^{2/3}}{(4n)^{1/3}} \cdot \frac{(2n^2)^{2/3}}{m^{2/3}} = n.$$

Consider a line $\ell \in \mathcal{L}$ that is defined by the equation $y = ax + b$, for some valid values of a and b. Notice that for any $x \in \{1, \ldots, r\}$, there exists a unique $y \in \{1, \ldots, 2rs\}$ such that the point (x,y) is incident to ℓ. That is, every line of $\mathcal{L}$ is incident to exactly r points of $\mathcal{P}$, and thus

$$I(\mathcal{P}, \mathcal{L}) = r \cdot |\mathcal{L}| = \frac{m^{2/3}}{(4n)^{1/3}} \cdot n = 2^{-2/3}m^{2/3}n^{2/3}.$$

$\square$

1.3   The crossing lemma

The *crossing number* of a graph $G = (V, E)$, denoted $\mathrm{cr}(G)$, is the smallest integer k such that we can draw G in the plane with k edge crossings. Figure 1.2(a) depicts a drawing of K_5 with a single crossing.[1] Since it is known that K_5 cannot be drawn without crossings, we have $\mathrm{cr}(K_5) = 1$. Given a graph $G = (V, E)$, we are interested

(a)   *(b)*
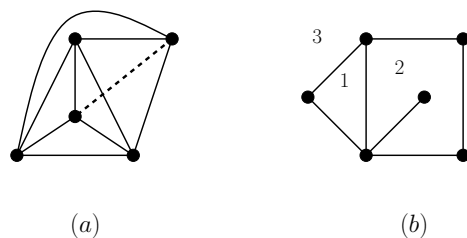
Figure 1.2: (a) A drawing of K_5 with a single crossing. (b) Two bounded faces and one unbounded.

in a lower bound for $\mathrm{cr}(G)$ with respect to $|V|$ and $|E|$. A graph G is said to be *planar* if $\mathrm{cr}(G) = 0$.

We consider a connected planar graph $G = (V, E)$ with v vertices and e edges. More specifically, we consider a drawing of G in the plane with no crossings. The *faces* of this drawing are the maximal two-dimensional connected regions that are bounded by the edges (including one outer, infinitely large region; e.g., see Figure 1.2(b)). Denote by f the number of faces in the drawing of G. According to *Euler's formula* (also known as Euler's polyhedron formula), we have

$$v + f = e + 2. \tag{1.1}$$

This formula does not hold for planar graphs that are not connected.

Every edge of G is either on the boundary of two faces or has both of its sides on the boundary of the same face. Moreover, the boundary of every face of G consists of at least three edges. Thus, we have $2e \geq 3f$. Plugging this into (1.1) yields

$$e = v + f - 2 \leq v + \frac{2e}{3} - 2.$$

That is, for any planar graph $G = (V, E)$, we have[2]

$$|E| \leq 3|V| - 6. \tag{1.2}$$

This inequality leads to our first lower bound on $\mathrm{cr}(G)$.

Lemma 1.4. *For any graph $G = (V, E)$, we have $\mathrm{cr}(G) \geq |E| - 3|V| + 6$.*

---

[1]The complete graph K_m has m vertices and an edge between every two vertices

[2]This is also valid for non-connected graphs, since the number of edges in Euler's formula becomes smaller when the graph is not connected.

Proof. Consider a drawing of G in the plane that minimizes the number of crossings. Let $E' \subset E$ be a maximum subset of the edges such that no two edges of E' intersect in the drawing. By (1.2), we have $|E'| \leq 3|V| - 6$. Since every edge of $E \setminus E'$ intersects at least one edge of E', and since $|E \setminus E'| \geq |E| - 3|V| + 6$, there are at least $|E| - 3|V| + 6$ crossings in the drawing. $\qquad\square$

Since K_5 has five vertices and ten edges, Lemma 1.4 yields the correct value $\mathrm{cr}(K_5) = 1$. However, in general the bound of this lemma is rather weak. For example, it is known that $\mathrm{cr}(K_n) = \Theta(n^4)$ while Lemma 1.4 implies only $\mathrm{cr}(K_n) = \Omega(n^2)$. We now amplify the lower bound of Lemma 1.4 by combining it with a probabilistic argument.

Lemma 1.5 (The crossing lemma). *Let $G = (V, E)$ be a graph with $|E| \geq 4|V|$. Then $\mathrm{cr}(G) = \Omega\left(|E|^3/|V|^2\right)$.*

Proof. Consider a drawing of G with $\mathrm{cr}(G)$ crossings. Set $p = \frac{4|V|}{|E|}$, and notice that by the assumption we have $0 < p \leq 1$. We remove every vertex of V from the drawing with probability $1 - p$ (together with the edges that are adjacent to it). Let $G' = (V', E')$ denote the resulting subgraph, and let c' denote the number of crossings that remain in the drawing.

To avoid confusion with the edge set E, we denote expectation of a random variable as $\mathbb{E}[\cdot]$. Since every vertex remains with probability p, we have $\mathbb{E}[|V'|] = p|V|$. Since every edge remains if and only if its two endpoints remain, we have $\mathbb{E}[|E'|] = p^2|E|$. Finally, since each crossing remains if and only if the two corresponding edges remain, we have $\mathbb{E}[c'] = p^4\mathrm{cr}(G)$. By linearity of expectation

$$
\begin{aligned}
\mathbb{E}[c' - |E'| + 3|V'|] &= p^4\mathrm{cr}(G) - p^2|E| + 3p|V| \\
&= \frac{4^4|V|^4}{|E|^4}\mathrm{cr}(G) - \frac{4^2|V|^2}{|E|^2} \cdot |E| + \frac{4|V|}{|E|} \cdot 3|V| \\
&= \frac{4^4|V|^4}{|E|^4}\mathrm{cr}(G) - \frac{4|V|^2}{|E|}.
\end{aligned}
$$

Since this is the expected value, there exists a subgraph $G^* = (V^*, E^*)$ with c^* crossings remaining from the drawing of G, such that

$$
c^* - |E^*| + 3|V^*| \leq \frac{4^4|V|^4}{|E|^4}\mathrm{cr}(G) - \frac{4|V|^2}{|E|}. \tag{1.3}
$$

By Lemma 1.4, we have $c^* \geq |E^*| - 3|V^*| + 6$. Combining this with (1.3) implies

$$0 < 6 \leq c^* - |E^*| + 3|V^*| \leq \frac{4^4 |V|^4}{|E|^4} \mathrm{cr}(G) - \frac{4|V|^2}{|E|}.$$

Tidying up this inequality yields the bound asserted in the lemma. $\quad\square$

It can be easily checked that the bound of Lemma 1.5 indeed implies $\mathrm{cr}(K_n) = \Omega(n^4)$. This lemma was originally derived in [1, 62].

1.4   Szemerédi-Trotter via the crossing lemma

We are now ready to prove Theorem 1.1. To help the reader, we first repeat the statement of the theorem.

Theorem 1.1. *Let $\mathcal{P}$ be a set of m points and let $\mathcal{L}$ be a set of n lines, both in $\mathbb{R}^2$. Then $I(\mathcal{P}, \mathcal{L}) = O(m^{2/3} n^{2/3} + m + n)$.*

Proof. We write $\mathcal{L} = \{\ell_1, \ldots, \ell_n\}$ and denote by m_i the number of points of $\mathcal{P}$ that are on ℓ_i. Notice that $I(\mathcal{P}, \mathcal{L}) = \sum_{i=1}^n m_i$. We may remove any line ℓ_i that satisfies $m_i = 0$, since this would have no effect on the number of incidences.

We build a graph $G = (V, E)$ as follows. Every vertex of V corresponds to a point of $\mathcal{P}$. For $v, u \in V$, we have $(v, u) \in E$ if v and u correspond to consecutive points along one of the lines of $\mathcal{L}$. Notice that ℓ_i corresponds to exactly $m_i - 1$ edges of E. Thus, we have $|V| = m$ and $|E| = \sum_{i=1}^n (m_i - 1) = I(\mathcal{P}, \mathcal{L}) - n$.

If $|E| < 4|V|$, then we immediately have $I(\mathcal{P}, \mathcal{L}) = O(m + n)$, as required. If $|E| \geq 4|V|$, then by Lemma 1.5 we have

$$\mathrm{cr}(G) = \Omega\left(\frac{(I(\mathcal{P}, \mathcal{L}) - n)^3}{m^2}\right). \tag{1.4}$$

We next draw G according to the point-line configuration — every vertex is at the corresponding point and every edge is the corresponding line segment. Since every crossing in this drawing corresponds to an intersection of two lines of $\mathcal{L}$, and since every two lines intersect at most once, we have $\mathrm{cr}(G) \leq \binom{n}{2} = O(n^2)$. Combining this with (1.4) implies

$$\frac{(I(\mathcal{P}, \mathcal{L}) - n)^3}{m^2} = O(n^2).$$

Rearranging this equation yields $I(\mathcal{P}, \mathcal{L}) = O(m^{2/3} n^{2/3} + n)$, as asserted. $\quad\square$

Notice that this proof is based on the double counting method. Specifically, to obtain a bound for the number of incidences, we counted $\text{cr}(G)$ in two different ways.

As in the proof of the weaker incidence bound in Lemma 1.2, we relied on the observation that two lines intersect at most once. This time we used a second geometric property, when stating that the line ℓ_i corresponds to exactly $m_i - 1$ edges of E. This step relied on the fact that a line consists of a single connected component, and does not intersect itself. When replacing the lines with other curves that satisfy the above properties, the proof of Theorem 1.1 remains valid.

1.5   The unit distances problem

The *unit distances problem* is one of the main open problems of Discrete Geometry. Although it has proven to be extremely difficult to solve, this problem is very easy to state: How many pairs of points in a planar set of n points could be at unit distance from each other? We denote the maximum number of such pairs as $u(n)$. By taking a set of n points equally spaced on a line, we immediately obtain $u(n) \geq n - 1$. Erdős [35] introduced the problem in 1946 and derived the bounds $u(n) = O(n^{3/2})$ and $u(n) = \Omega(n^{1+c/\log\log n})$ (for some constant c). Even though this is such a central problem in Discrete Geometry, in the seven decades that have passed the lower bound was never improved and the upper bound was improved only once. The bound $u(n) = O(n^{4/3})$ was derived by Spencer, Szemerédi, and Trotter [92] in 1984.

Consider a set $\mathcal{P} \subset \mathbb{R}^2$ of n points such that the number of unit distances between pairs of points of $\mathcal{P}$ is $u(n)$. We draw a unit circle (a circle of radius one) around each point of $\mathcal{P}$, and denote the set of these n circles as $\mathcal{C}$. Every two points $p, q \in \mathcal{P}$ that determine a unit distance correspond to two incidences in $\mathcal{P} \times \mathcal{C}$ — the circle around p is incident to q and vice versa. Thus, to bound $u(n)$ it suffices to bound the maximum number of incidences between n points and n unit circles (it is not hard to show that the two expressions are in fact asymptotically equivalent).

Theorem 1.6. *Let $\mathcal{P}$ be a set of n points and let $\mathcal{C}$ be a set of n unit circles, both in $\mathbb{R}^2$. Then $I(\mathcal{P}, \mathcal{C}) = O(n^{4/3})$.*

Notice that the theorem immediately implies the current best bound $u(n) = O(n^{4/3})$.

Proof. We imitate the proof of Theorem 1.1. Let $\mathcal{C} = \{c_1, \ldots, c_n\}$ and denote by m_i the number of points of $\mathcal{P}$ that are on c_i. Notice that $I(\mathcal{P}, \mathcal{C}) = \sum_{i=1}^{n} m_i$. We may remove any circle c_i that satisfies $m_i < 3$, since these circles yield at most $2n$ incidences.

We build a graph $G = (V, E)$ as follows. Every vertex of V corresponds to a point of $\mathcal{P}$. For $v, u \in V$, we have $(v, u) \in E$ if v and u are consecutive points along at least one of the circles of $\mathcal{C}$. Notice that c_i corresponds to exactly m_i edges of E, and that every edge originates from at most two unit circles. Thus, we have $|V| = n$ and $|E| \geq \sum_{i=1}^{n} m_i/2 = I(\mathcal{P}, \mathcal{C})/2$.

If $|E| < 4|V|$, then we immediately have $I(\mathcal{P}, \mathcal{C}) = O(n)$, as required. If $|E| \geq 4|V|$, then by Lemma 1.5 we have

$$\mathrm{cr}(G) = \Omega\left(\frac{I(\mathcal{P}, \mathcal{C})^3}{n^2}\right). \tag{1.5}$$

We next draw G according to the point-circle configuration — every vertex is at the corresponding point and every edge is the corresponding circle arc. Since every crossing in the drawing corresponds to an intersection of two circles of $\mathcal{C}$, and since every two circles intersect at most twice, we have $\mathrm{cr}(G) \leq 2\binom{n}{2} = O(n^2)$. Combining this with (1.5) implies

$$\frac{I(\mathcal{P}, \mathcal{C})^3}{n^2} = O(n^2).$$

Rearranging this equation yields $I(\mathcal{P}, \mathcal{C}) = O(n^{4/3})$, as asserted. $\qquad\square$

The common belief seems to be that the following conjecture holds.

Conjecture 1.7. $u(n) = O(n^{1+\varepsilon})$ *for any* $\varepsilon > 0$.

The above is a good example for how little we currently know about incidences. Even though the case of incidences with lines in $\mathbb{R}^2$ has been settled for decades, already when moving to unit circles the problem is wide open.

As another indication that the unit distances problem is deeper than it might at first seem, the answer to this problem significantly depends on the norm that is used:

- For ℓ_2 this is a long-standing difficult problem.
- For ℓ_1 and ℓ_∞ the problem is trivially $\Theta(n^2)$.
- Valtr [102] showed that there is a simply defined norm for which the answer is $\Theta(n^{4/3})$.
- Matoušek [65] showed that for a generic norm the maximum number of unit distances is $O(n \log n \log \log n)$.

Note that the conjectured bound $n^{1+c/\log\log n}$ for ℓ_2 is different from all of the bounds stated above. This is why one may say that this is a problem about the underlying geometry. In particular, about studying properties that are unique for the Euclidean distance.

1.6   The distinct distances problem

The *distinct distances* problem can be considered as the twin problem of the unit distances problem, and it was introduced in the same 1946 paper of Erdős [35]. The question asks for the minimum number of distinct distances that can be determined by a set of n points in the plane. That is, denoting the distance between two points $p, q \in \mathbb{R}^2$ as $|pq|$, we wish to find $\min_{|\mathcal{P}|=n} |\{|pq| : p, q \in \mathcal{P}\}|$. We denote this quantity as $d(n)$.

It can be easily verified that a set of n points that are equally spaced on a line determines $n - 1$ distinct distances. Thus, we have $d(n) \le n - 1$. A better bound appeared in Erdős' original paper. Specifically, Erdős considered a $\sqrt{n} \times \sqrt{n}$ integer lattice. The number of distances that are determined by this set is an immediate corollary of a result from number theory.

Theorem 1.8. (Landau and Ramanujan [11, 60]) *The number of positive integers smaller than n that are the sum of two squares is $\Theta(n/\sqrt{\log n})$.*

Every distance in the $\sqrt{n} \times \sqrt{n}$ integer lattice is the square root of a sum of two squares between 0 and n. Thus, Theorem 1.8 implies that the number of distinct distances in this case is $\Theta(n/\sqrt{\log n})$.

Theorem 1.9 (Erdős [35]). $d(n) = O(n/\sqrt{\log n})$.

For the lower bound on $d(n)$, we begin by deriving Erdős' original bound (using a different proof).

Claim 1.10. $d(n) = \Omega(n^{1/2})$.

Proof. Consider an n point set $\mathcal{P}$ and two points $v, u \in \mathcal{P}$. Let d_v denote the number of distinct distances between v and $\mathcal{P} \setminus \{v\}$. Notice that the points of $\mathcal{P} \setminus \{v\}$ are contained in d_v circles that are centered at v. We denote this set of circles as $\mathcal{C}_v$. We define d_u and $\mathcal{C}_u$ symmetrically. Each of the $n - 2$ points of $\mathcal{P} \setminus \{v, u\}$ is contained in the intersection of a circle of $\mathcal{C}_v$ and a circle of $\mathcal{C}_u$. Since the number of such intersections is at most $2|\mathcal{C}_v||\mathcal{C}_u| = 2d_v d_u$, we have $2d_v d_u \ge n - 2$, which in turn implies $\max\{d_v, d_u\} = \Omega(n^{1/2})$. (An example is depicted in Figure 1.3.) $\square$

We now derive an improved bound by using incidences. This bound was originally derived by Moser [68] in 1952, using a different argument.

Claim 1.11. $d(n) = \Omega(n^{2/3})$.
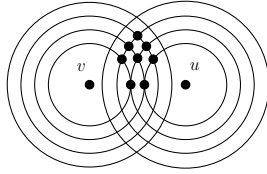
Figure 1.3: The points of $\mathcal{P} \setminus \{v, u\}$ are contained in the intersections of $\mathcal{C}_v$ and $\mathcal{C}_u$.

Proof. Consider an n point set $\mathcal{P}$ that determines d distinct distances, and denote these distances as $D = \{\delta_1, \ldots, \delta_d\}$. Let $\mathcal{C}$ denote the set of $n \cdot d$ circles with a center in $\mathcal{P}$ and a radius in D. The claim is proved by double counting $I(\mathcal{P}, \mathcal{C})$.

For every point $v \in \mathcal{P}$, the points of $\mathcal{P} \setminus \{v\}$ are contained in the d circles of $\mathcal{C}$ centered at v. Thus, $I(\mathcal{P}, \mathcal{C}) = n(n-1)$.

Next, let $\mathcal{C}_i$ denote the subset of circles of $\mathcal{C}$ with radius δ_i. By Theorem 1.6, we have $I(\mathcal{P}, \mathcal{C}_i) = O(n^{4/3})$ (notice that Theorem 1.6 is valid for any set of circles with the same radii). Thus, we have

$$I(\mathcal{P}, \mathcal{C}) = \sum_{i=1}^{d} I(\mathcal{P}, \mathcal{C}_i) = O(dn^{4/3}).$$

Combining our two bounds for $I(\mathcal{P}, \mathcal{C})$ immediately implies the assertion of the claim. $\qquad \square$

A simpler proof of Claim 1.11 goes as follows. Each of the $\binom{n}{2}$ pairs of points determines a distance. By Theorem 1.6 any distance occurs $O(n^{4/3})$ times, so to cover $\Theta(n^2)$ pairs there must be $\Omega(n^{2/3})$ distinct distances. We presented the longer proof since it sheds more light about how to use incidences.

Both proofs show that the distinct distances problem can in some sense be reduced to the unit distances problem. An upper bound of $u(n) = O(n^{1+c/\log\log n})$ would yield an almost tight bound for $d(n)$.

In 2010, Guth and Katz [51] proved the almost tight bound $d(n) = \Omega(n/\log n)$. Unlike the two proofs above, this is a deep result that combines tools from several different fields. One of the peaks of this book is a proof of this result. Even though the distinct distances problem is solved (up to a gap of $\sqrt{\log n}$), interesting variants of it are still wide open. A couple of examples:

- The problem is still open in $\mathbb{R}^d$ for any $d \geq 3$. Erdős constructed a set of n points in $\mathbb{R}^d$ that determines $\Theta(n^{2/d})$ distinct distances, and conjectured that no set determines a smaller number. So far no one managed to apply the polynomial method even for the case of $\mathbb{R}^3$.

- In $\mathbb{R}^2$, characterizing the n point sets that determine $O(n/\log n)$ distinct distances seems to be a very difficult problem. The past several decades yielded many conjectures regarding this but hardly any results.

For a list of many other related open problems, see [85].

1.7   A problem about unit area triangles

In this section we briefly mention a problem that can be considered as one of the many generalizations of the unit distances problem. The problem is: What is the maximum number of unit area triangles that have their vertices in a set of n points in $\mathbb{R}^2$?

Consider two points $p, q \in \mathbb{R}^2$ at a distance of d from each other. A key observation is that p and q form a unit area triangle with a third point r if and only if r is on one of the two lines that are parallel to the segment pq and at a distance of $2/d$ from this segment (e.g., see Figure 1.4(a)). Thus, by taking two parallel lines at a distance of 2 from each other, and placing $n/2$ points at unit intervals on each, we obtain $\Theta(n^2)$ unit triangles (e.g., see Figure 1.4(b)). Erdős and Purdy [39] showed that a $\sqrt{\log n} \times n/\sqrt{\log n}$ section of the integer lattice determines $\Omega(n^2 \log \log n)$ triangles of the same area.

(a)   (b)
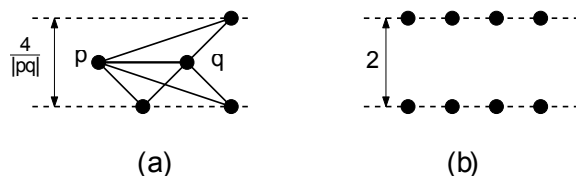
Figure 1.4: (a) The points that form a unit triangle with p and q are on two parallel lines. (b) A configuration with $\Theta(n^2)$ unit triangles.

Claim 1.12 (Pach and Sharir [69]). *Every planar set of n points determines $O(n^{7/3})$ unit triangles.*

Proof. Consider a set $\mathcal{P}$ of n points. For a point $p \in \mathcal{P}$ we bound the number of unit triangles that are determined by p and two other points of $\mathcal{P}$. For any $q \in \mathcal{P} \setminus \{p\}$, we denote by ℓ_{pq}, ℓ'_{pq} the lines that are parallel to the segment pq and at a distance of $2/|pq|$ from it. We set $\mathcal{L}_p = \{\ell_{pq}, \ell'_{pq} : p, q \in \mathcal{P}\}$. Notice that any line $\ell_{p,q}$ can originate from at most two points $q \in \mathcal{P} \setminus \{p\}$. Thus, $n - 1 \le |\mathcal{L}_p| \le 2n - 2$. The number of unit triangles that involve p is at least $I(\mathcal{P}, \mathcal{L}_p)/2$. By Theorem 1.1, we

have $I(\mathcal{P}, \mathcal{L}_p) = O(n^{4/3})$. The assertion of the claim is obtained by summing this bound over every $p \in \mathcal{P}$. $\qquad\square$

Recently, Raz and Sharir [74] improved this bound to $O(n^{20/9})$ by considering incidences with two-dimensional surfaces in $\mathbb{R}^4$.

1.8 The sum-product problem

In this section we examine an application of the Szemerédi-Trotter theorem for a problem that at first may not seem related to geometry. Given a set A of n real numbers, we consider the sets

$$A + A = \{a + b : a, b \in A\}, \quad \text{and} \quad AA = \{ab \mid a, b \in A\}.$$

It is not difficult to find sets A that satisfy $|A + A| = \Theta(n)$. For example, we can take A to be $\{1, 2, 3, \ldots, n\}$, or any other arithmetic progression. Similarly, to obtain $|AA| = \Theta(n)$, we can take A to be a geometric progression. Erdős and Szemerédi [40] made the following conjecture.

Conjecture 1.13. *For any $\varepsilon > 0$, there exists n_0, such that any set A of $n > n_0$ integers satisfies*

$$\max\{|A + A|, |AA|\} = \Omega(n^{2-\varepsilon}).$$

Over the years this question has been generalized to various fields, and received the name *the sum-product problem.* In 1997, Elekes [30] introduced a geometric approach for the sum-product problem, which influenced many later works. We now study Elekes' result.

Theorem 1.14. *Let A be a set of n real numbers. Then*

$$\max\{|A + A|, |AA|\} = \Omega(n^{5/4}).$$

Proof. Consider the point set

$$\mathcal{P} = \{(c, d) \ : \ c \in A + A \quad \text{and} \quad d \in AA\}.$$

Notice that $|\mathcal{P}| = |A + A| \cdot |AA|$. We also consider the set of lines

$$\mathcal{L} = \{y = a(x - a') \ : \ a, a' \in A\}$$

(where by $y = a(x - a')$ we refer to the line that is defined by this expression). Notice that $|\mathcal{L}| = n^2$.

The proof is based on double counting $I(\mathcal{P}, \mathcal{L})$. First, a line that is defined by the equation $y = a(x - a')$ (with $a, a' \in A$) contains the points of $\mathcal{P}$ of the form $(a' + b, ab)$ for every $b \in A$. That is, every line of $\mathcal{L}$ is incident to at least n points of $\mathcal{P}$. Therefore, we have

$$I(\mathcal{P}, \mathcal{L}) \geq |\mathcal{L}|n = n^3.$$

On the other hand, by applying Theorem 1.1 we obtain

$$I(\mathcal{P}, \mathcal{L}) = O\left(|\mathcal{P}|^{2/3}|\mathcal{L}|^{2/3} + |\mathcal{P}| + |\mathcal{L}|\right)$$
$$= O\left(|A + A|^{2/3}|AA|^{2/3}n^{4/3} + |A + A| \cdot |AA| + n^2\right).$$

By combining the two bounds for $I(\mathcal{P}, \mathcal{L})$, we get

$$n^3 = O\left(|A + A|^{2/3}|AA|^{2/3}n^{4/3} + |A + A| \cdot |AA| + n^2\right),$$

or

$$|A + A| \cdot |AA| = \Omega\left(n^{5/2}\right),$$

which implies the assertion of the theorem. □

Elekes' bound has been improved several times, always using geometric arguments. Most notably, Solymosi [89] derived the bound $\Omega(n^{4/3}/\log^{1/3} n)$.
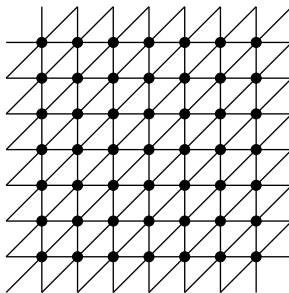
Figure 1.5: A construction with $n/4$ vertical lines, $n/4$ horizontal lines, $n/2$ diagonal lines, and $\Theta(n^2)$ points that are 3-rich.

1.9   Rich points

Given a set of n lines in $\mathbb{R}^2$ and an integer $r \geq 2$, we say that a point $p \in \mathbb{R}^2$ is *r-rich* if at least r lines of $\mathcal{L}$ are incident to p. Note that a 4-rich point is also 3-rich and also 2-rich. In a set of n lines that intersect at the origin there is a single 2-rich point — the origin. Let $M_{\geq r}(n)$ denote the maximum number of r-rich points that a set of n lines in $\mathbb{R}^2$ can have. For example, Figure 1.5 demonstrates that $M_{\geq 3}(n) = \Omega(n^2)$. Rich points have an important role in advanced proofs that we will see in Chapter 9.

Rich points lead us to an equivalent formulation of the Szemerédi-Trotter theorem. By 'equivalent', we mean that each formulation can be easily derived from the other by using only basic arguments.

Lemma 1.15. *The Szemerédi-Trotter theorem is equivalent to the claim that $M_{\geq r}(n) = O\left(\frac{n^2}{r^3} + \frac{n}{r}\right)$ holds for every $r \geq 2$.*

Proof. We first prove that Szemerédi-Trotter bound implies $M_{\geq r}(n) = O(n^2/r^3)$ for every $r \geq 2$. We have $M_{\geq r}(n) = O(n^2)$ for every $r \geq 2$, since there are $O(n^2)$ intersection points in a set of n lines. This completes the proof when r is a constant, so we may assume that r is larger than the constant in the $O(\cdot)$-notation of the Szemerédi-Trotter bound.

Consider a set $\mathcal{L}$ of n lines in $\mathbb{R}^2$ and a sufficiently large value of r. Let $\mathcal{P}$ denote the set of points that are incident to at least r lines of $\mathcal{L}$, and set $m_r = |\mathcal{P}|$. By definition, we have $I(\mathcal{P}, \mathcal{L}) \geq m_r r$. On the other hand, the Szemerédi-Trotter bound implies $I(\mathcal{P}, \mathcal{L}) = O(m_r^{2/3} n^{2/3} + n + m_r)$. Combining these two bounds yields $m_r r = O(m_r^{2/3} n^{2/3} + n + m_r)$. Since r is larger than the constant in the $O(\cdot)$-notation, it cannot be that the dominating term inside of the $O(\cdot)$-notation is m_r, and thus $m_r r = O(m_r^{2/3} n^{2/3} + n)$. This immediately implies $m_r = O\left(\frac{n^2}{r^3} + \frac{n}{r}\right)$, as required.

We now assume that $M_{\geq r}(n) = O(n^2/r^3 + n/r)$ holds for every $r \geq 2$, and rely on this to prove the Szemerédi-Trotter bound. Consider a set $\mathcal{P}$ of m points and a set $\mathcal{L}$ of n lines. If $m = \Omega(n^2)$, then by Lemma 1.2 we have $I(\mathcal{P}, \mathcal{L}) = O(n\sqrt{m} + m) = O(m)$. Thus, we may assume that $m = O(n^2)$.

Let $\hat{m}_i$ denote the number of points of $\mathcal{P}$ that are incident to more than 2^{i-1} lines of $\mathcal{L}$ and to at most 2^i such lines. Let $\mathcal{P}_+$ be the set of points of $\mathcal{P}$ that are incident to more than $\sqrt{n}$ lines of $\mathcal{L}$, and set $k = \left\lceil \log\left(n^{2/3}/m^{1/3}\right)\right\rceil$. Since $m = O(n^2)$, we get that $k \geq 1$. Thus,

$$I(\mathcal{P}, \mathcal{L}) \leq \sum_{i \geq 0}^{(\log n)/2} \hat{m}_i 2^i + I(\mathcal{P}_+, \mathcal{L}) = \sum_{i=0}^{k} \hat{m}_i 2^i + \sum_{i=k+1}^{(\log n)/2} \hat{m}_i 2^i + I(\mathcal{P}_+, \mathcal{L}). \qquad (1.6)$$

If $k \geq \frac{1}{2}\log n$, we ignore the second sum in (1.6) and have the index of the first sum stop at $\frac{1}{2}\log n$. Since $\hat{m}_i \leq m$ obviously holds for every i, we have

$$\sum_{i=0}^{k} \hat{m}_i 2^i \leq \sum_{i=0}^{k} m 2^i = O\left(m^{2/3}n^{2/3} + m\right).$$

For the second sum of (1.6), we notice that when $i \leq \sqrt{n}$ the bound on $M_{\geq r}(n)$ yields $\hat{m}_i = O(n^2/2^{3i})$. Thus, we get

$$\sum_{i=k+1}^{(\log n)/2} \hat{m}_i 2^i = \sum_{i=k+1}^{(\log n)/2} O\left(\frac{n^2}{2^{2i}}\right) = O\left(m^{2/3}n^{2/3}\right).$$

It remains to bound $I(\mathcal{P}_+, \mathcal{L})$. By the bound on $M_{\geq r}(n)$, we have $|\mathcal{P}_+| = O(\sqrt{n})$. By plugging this into the bound $I(\mathcal{P}, \mathcal{L}) = O(m\sqrt{n} + n)$ from Lemma 1.2, we obtain $I(\mathcal{P}_+, \mathcal{L}) = O(n)$.   □

Given a set of m points in $\mathbb{R}^2$ and an integer r, we say that a line $\ell \subset \mathbb{R}^2$ is r-rich if ℓ is incident to at least r of the points. This leads to yet another equivalent formulation of the Szemerédi-Trotter theorem (see Problem 1.15).

1.10   Exercises

Problem 1.1. Let $\mathcal{P}$ be a set of m points in $\mathbb{R}^2$. Derive an upper bound on the number of lines that contain at least k points of $\mathcal{P}$, for any $k \geq 2$.

Problem 1.2. Construct a set $\mathcal{P}$ of m points and a set Γ of n parabolas that are defined by equations of the form $y = ax^2 + bx + c$, such that $I(\mathcal{P}, \Gamma) = \Theta(m^{1/2}n^{5/6})$ (hint: Adapt the proof of Claim 1.3).

Problem 1.3. Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^2$ and let Γ be a set of n distinct hyperbolas, where each hyperbola is defined by an equation $(x - a)^2 - (y - b)^2 = 1$ (for some $a, b \in \mathbb{R}$). Use the crossing lemma to prove that $I(\mathcal{P}, \Gamma) = O(n^{4/3})$.

Problem 1.4. Prove that the maximum number of right-angled triangles that can be determined by a set of n points in $\mathbb{R}^2$ is $O\left(n^{7/3}\right)$ (be careful about having identical lines).

Problem 1.5. Define a *unit chain* as a triple of points $(a, b, c) \in \mathbb{R}^2$ such that $|ab| = |bc| = 1$ and $a \neq c$. Derive asymptotically tight bounds for the maximum

number of unit chains that a set of n points in $\mathbb{R}^2$ can span. (While the unit distances problem is extremely difficult, this variant is surprisingly simple).

Problem 1.6. Given a set $\mathcal{P}$ of n points in $\mathbb{R}^2 \setminus \{(0,0)\}$, we consider the set of triangles whose vertices are the origin and two points from $\mathcal{P}$. If two such triangles are congruent, we say that they belong to the same congruency class. Derive asymptotically tight bounds for the minimum number of triangle congruency classes that $\mathcal{P}$ can span.

Problem 1.7. ([99]) Let A be a set of n real numbers. We define

$$A + AA = \{a + bc : a, b, c \in A\}.$$

Prove that $|A + AA| = \Omega(n^{3/2})$. (Hint: Consider the point set $A \times (A + AA)$.)

Problem 1.8. Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^2$. Prove that for at least $n-1$ points $p \in \mathcal{P}$ there are $\Omega(n^{1/2})$ distinct distances between p and the points of $\mathcal{P} \setminus \{p\}$.

Problem 1.9. ([9]) Let $\mathcal{P}$ be a set of m points in $\mathbb{R}^2$, and let $\mathcal{L}$ denote the set of all lines that are incident to at least two points of $\mathcal{P}$. Prove that either there exists a line containing $\Omega(m)$ points of $\mathcal{P}$ or $|\mathcal{L}| = \Omega(m^2)$.

To prove the claim, set $\mathcal{L}_j = \{\ell \in \mathcal{L} : 2^j \leq |\mathcal{P} \cap \ell| < 2^{j+1}\}$. There are $\Theta(n^2)$ pairs of points of $\mathcal{P}$, and each pair is contained in one line of $\mathcal{L}$. Prove that there exists a sufficiently large constant c, such that at most $m^2/100$ pairs are on the lines of $\bigcup_{j=\log c}^{\log(m/c)} \mathcal{L}_j$. Then consider the remaining pairs.

Problem 1.10. Let A be a set of m real numbers. Prove that the number of collinear triples of points[3] in the lattice $A \times A \subset \mathbb{R}^2$ is $O(m^4 \log m)$.

Problem 1.11. In Section 1.5 we mentioned that the behaviour of the unit distances problem depends on the notion of distance that is used. The L_1 distance (or *Manhattan distance*) between two points $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ is $|x_1 - x_2| + |y_1 - y_2|$. Find an asymptotically tight bound for the unit distances problem when using the L_1 distance.

Problem 1.12. Find an asymptotically tight bound for the distinct distances problem when using the L_1 distance (as defined in the previous problem).

Problem 1.13. The following theorem is taken from [52].

> **Theorem.** *Let $\mathcal{L}$ be a set of lines and let $\mathcal{P}$ be a set of m points, both in $\mathbb{R}^3$, such that each line of $\mathcal{L}$ contains at least r points of $\mathcal{P}$. If $|\mathcal{L}| = \Omega\left(m^2/r^4 + m/r\right)$*

---

[3]three points are collinear if there exists a line that is incident to all three.

then there exists a plane containing $\Omega(m/r^2)$ points of $\mathcal{P}$.

Rely on this theorem to prove the following corollary, while also finding what the question marks should be replaced with.

> **Corollary.** *Let $\mathcal{L}$ be a set of n lines and let $\mathcal{P}$ be a set of m points, both in $\mathbb{R}^3$, such that every plane contains $O(???)$ points of $\mathcal{P}$. Then $I(\mathcal{P}, \mathcal{L}) = O\left(m^{1/2}n^{3/4} + m\log m + n\right)$.*

Problem 1.14. ([41]) A matrix is said to be *totally positive* if all of its minors are positive. Let M be an $n \times 2$ totally positive matrix. Prove that the number of 2×2 minors of M that are equal to 1 is $O(n^{4/3})$.

Problem 1.15. Adapt the proof of Theorem 1.15 to show that the Szemerédi-Trotter theorem is equivalent to a bound on the number of r-rich lines in $\mathbb{R}^2$.

1.11   Open problems

In this chapter we focused mainly on point-line incidences in $\mathbb{R}^2$, which is one of the few incidence problems that are completely settled. In general, after obtaining an asymptotically tight bound for an extremal combinatorics problem, the next step is to characterize the configurations that achieve this bound. In Discrete Geometry, problems of characterizing the extremal configurations tend to be unusually difficult, and only few such problems are solved. The case of point-line incidences in $\mathbb{R}^2$ is no different, in the sense that not much is known about point-line configurations that determine a large number of incidences. In Claim 1.3, we saw Elekes' construction for obtaining the Szemerédi–Trotter bound $\Theta(m^{2/3}n^{2/3})$. Both this construction and Erdős' earlier construction consider a point set that is a rectangular section of the integer lattice. One can obtain somewhat different point sets by applying various projective transformations on these constructions.

Conjecture 1.16. *Consider sufficiently large positive integers m and n that satisfy $m = O(n^2)$ and $m = \Omega(\sqrt{n})$. Let $\mathcal{P}$ be a set of m points and $\mathcal{L}$ be a set of n lines, both in $\mathbb{R}^2$, such that $I(\mathcal{P}, \mathcal{L}) = \Theta(m^{2/3}n^{2/3})$. Then there exists a subset $\mathcal{P}' \subset \mathcal{P}$ such that $|\mathcal{P}'| = \Theta(m)$ and $\mathcal{P}'$ is contained in a section of the integer lattice of size $\Theta(m)$, possibly after applying a projective transformation to it.*

Note that we already mentioned several open problems throughout the chapter: the unit distances problem, two distinct distances problems, the maximum number of unit area triangles, and the sum-product problem.

Chapter 2

Basic Real Algebraic Geometry in $\mathbb{R}^2$

> *"Everyone knows what a curve is, until he has studied enough mathematics to become confused through the countless number of possible exceptions"* / attributed to Felix Klein [16].

This chapter is a very basic introduction to algebraic geometry over the reals. At this point we focus mainly on the plane $\mathbb{R}^2$, postponing the treatment of $\mathbb{R}^d$ to Chapter 4. This allows us to discuss several planar results in Chapter 3, before dealing with more involved algebraic geometry.

2.1   Varieties

Algebraic geometry can be thought of as the study of geometries that arise from algebra (or more specifically, from polynomials). In this section we present varieties, which are central geometric objects of Algebraic Geometry.

The *polynomial ring* $\mathbb{R}[x_1, \ldots, x_d]$ is the set of polynomials in the variables $x_1, \ldots, x_d$ and with coefficients in $\mathbb{R}$. Given a (possibly infinite) set of polynomials $f_1, \ldots, f_k \in \mathbb{R}[x_1, \ldots, x_d]$, the *affine variety* $\mathbf{V}(f_1, \ldots, f_k)$ is defined as

$$\mathbf{V}(f_1, \ldots, f_k) = \{(a_1, \ldots, a_d) \in \mathbb{R}^d : f_j(a_1, \ldots, a_d) = 0 \text{ for all } 1 \le j \le k\}.$$

The adjective "affine" distinguishes the variety from projective varieties. At this point we only consider affine varieties, and for brevity refer to those simply as varieties.[1] For example, some varieties in $\mathbb{R}^3$ are a torus, a union of a circle and a line,

---

[1] Some authors call these objects *algebraic sets*, while using the word variety for what we will refer to as an irreducible variety.

and a set of 1000 points.

The following is a special case of *Hilbert's basis theorem* (e.g., see [23, Section 2.5]).

Theorem 2.1. *Every variety can be described by a finite set of polynomials.*

Theorem 2.1 is valid in every field. When working over the reals, we can say something stronger.

Corollary 2.2. *Every variety in* $\mathbb{R}^d$ *can be described by a single polynomial.*

Proof. Consider a variety $U \subset \mathbb{R}^d$. By Theorem 2.1, there exist $f_1, \ldots, f_k \in \mathbb{R}[x_1, \ldots, x_d]$ such that $U = \mathbf{V}(f_1, \ldots, f_k)$. We set $f = f_1^2 + f_2^2 + \cdots + f_k^2$. Notice that for any point $p \in \mathbb{R}^d$ we have $f(p) = 0$ if and only if $f_1(p) = \cdots = f_k(p) = 0$. Thus, we have $U = \mathbf{V}(f)$. $\qquad\square$

We consider some basic properties of varieties.

Claim 2.3. *Let* $U, W \subset \mathbb{R}^d$ *be two varieties, and let* $\tau : \mathbb{R}^d \to \mathbb{R}^d$ *be an invertible linear map (such as a translation, rotation, reflection, or stretching). Then*
(a) $U \cap W$ *is a variety,*
(b) $U \cup W$ *is a variety, and*
(c) $\tau(U)$ *is a variety.*

Proof. Since U and W are varieties, there exist $f_1, \ldots, f_k, g_1, \ldots, g_m \in \mathbb{R}[x_1, \ldots, x_d]$ such that $U = \mathbf{V}(f_1, \ldots, f_k)$ and $W = \mathbf{V}(g_1, \ldots, g_m)$.[2] For (a), notice that we have $U \cap W = \mathbf{V}(f_1, \ldots, f_k, g_1, \ldots, g_m)$. For (b), we have $U \cup W = \mathbf{V}(H)$, where

$$H = \bigcup_{\substack{1 \le i \le k \\ 1 \le j \le m}} \{f_i \cdot g_j\}.$$

For (c), we write the inverse of τ as $\psi \in (\mathbb{R}[x_1, \ldots, x_d])^d$. Then

$$\tau(U) = \mathbf{V}\left(f_1 \circ \psi, \ldots, f_d \circ \psi\right). \qquad\square$$

At this point it might be instructive to ask what subsets of $\mathbb{R}^d$ are not varieties.

Claim 2.4. *The set* $X = \{(x, x) : x \in \mathbb{R}, \ x \ne 1\} \subset \mathbb{R}^2$ *is not a variety.*

---

[2]By Corollary 2.2, it suffices to use a single polynomial for each variety. We present this slightly less elegant proof since it applies in every field.

Proof. Assume for contradiction that there exist $f_1, \ldots, f_k \in \mathbb{R}[x_1, x_2]$ such that $X = \mathbf{V}(f_1, \ldots, f_k)$. For every $1 \le i \le k$, we set $g_i(t) = f_i(t, t)$ and note that $g_i \in \mathbb{R}[t]$. Since $g_i(t)$ vanishes on every $t \ne 1$ we have that $g_i(t) = 0$ (recall that 0 is the only univariate polynomial that has infinitely many zeros). This in turn implies that $f_i(1, 1) = 0$. Since this holds for every $1 \le i \le k$, we get a contradiction to $(1, 1) \notin X$. $\qquad\square$

Similarly, a line segment and half a circle are not varieties. For other types of sets that are not varieties, see Problem 2.1.

We say that a set U' is a *subvariety* of a variety U if $U' \subseteq U$ and U' is a variety. We say that U' is a *proper subvariety* of U if U' is non-empty a subvariety of U' and $U' \ne U$. A variety U is *reducible* if there exist two proper subvarieties $U', U'' \subset U$ such that $U = U' \cup U''$. Otherwise, U is *irreducible*. For example, the union of the two axes $\mathbf{V}(xy) \subset \mathbb{R}^2$ is reducible since $\mathbf{V}(xy) = \mathbf{V}(x) \bigcup \mathbf{V}(y)$.

Every variety U can be decomposed into distinct irreducible subvarieties $U_1, U_2, \ldots, U_k$ such that $U = \bigcup_{i=1}^{k} U_i$ (to see why the number of such components is finite, see Problem 4.1). After removing every U_i that is a proper subvariety of another U_j, we obtain a unique decomposition of U. The subvarieties of this decomposition are said to be the *irreducible components of U* (or *components*, for brevity).

2.2 Curves in $\mathbb{R}^2$

Chapter 4 contains a detailed discussion about degrees, singular points, and other basic properties of varieties in $\mathbb{R}^d$. At this point we only consider the case of $\mathbb{R}^2$, where these concepts are significantly simpler. For now we also avoid defining the dimension of a variety.

We say that an irreducible variety in $\mathbb{R}^2$ is a *curve* if it is not a single point or one of the trivial varieties $\emptyset$ and $\mathbb{R}^2$ (note that a set of several points is reducible). A reducible variety is a curve if each of its components is a curve. This definition corresponds to what we would intuitively call a polynomial curve.

Degrees and intersections. We say that the *degree* of a curve $\gamma \subset \mathbb{R}^2$ is the minimum integer k such that there exists a polynomial $f \in \mathbb{R}[x, y]$ of degree k with $\mathbf{V}(f) = \gamma$.

We now present a real variant of *Bézout's theorem*. This theorem is ubiquitous in this book, with a large variety of applications. Proving the theorem requires more advanced algebraic geometry tools, and is postponed to Chapter ???.

Theorem 2.5 (Bézout's theorem). *Let f and g be two polynomials in $\mathbb{R}[x,y]$ of degrees k_f and k_g, respectively. If f and g do not have common factors, then $\mathbf{V}(f) \cap \mathbf{V}(g)$ consists of at most $k_f \cdot k_g$ points.*

As simple examples, notice that two lines indeed intersect in at most one point and that two ellipses (which are of degree 2) intersect in at most four points.
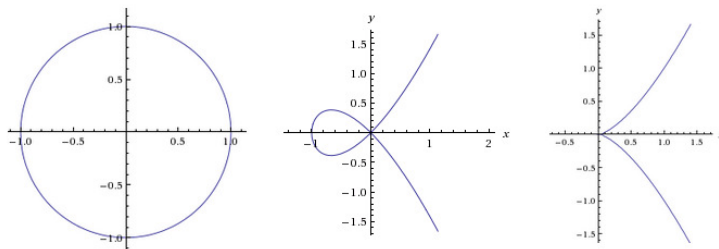
Figure 2.1: Every point of the circle has a well defined tangent line. In the other two curves the tangent is not well defined at the origin. These curves are $\mathbf{V}(y^2 - x^3 - x^2)$ and $\mathbf{V}(x^3 - y^2)$.

Singular points Consider a curve $\gamma \subset \mathbb{R}^2$ and a point $p \in \gamma$. Intuitively (and with some exceptions), p is a *singular point* of γ if one of the following holds:

- The tangent line to γ at p is not well defined. For example, see Figure 2.1.
- The point p is contained in more than one irreducible component of γ. For example, consider the union of two circles that intersect at a point p and have the same tangent line at p. Although one might say that the tangent line is well defined at p, we still consider p as a singular point.
- The point p is an *isolated point* of γ. That is, there exists an open set that contains p and no other point of γ.

A point of γ that is not singular is said to be a *regular* point of γ. For a reader that is unfamiliar with real algebraic geometry, it might seem as if the third bullet is redundant, since a curve cannot have a component that is a single point. To see why this is not the case, consider the cubic curve $\mathbf{V}(y^2 - x^3 + x^2)$ which is depicted in Figure 2.2. Even though this is an irreducible variety, it consists of a connected component that looks like a curve together with the origin. The set obtained by removing the origin is not a variety. Thus, by the third bullet, the origin is a singular point of the curve $\mathbf{V}(y^2 - x^3 + x^2)$.

We now provide a rigorous definition of a singular point of a curve $\gamma \subset \mathbb{R}^2$. Given a polynomial $f \in \mathbb{R}[x,y]$, the *gradient* of f is

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right).$$

Figure 2.2: The irreducible curve $\mathbf{V}(y^2 - x^3 + x^2)$.

Consider a minimum-degree polynomial $f \in \mathbb{R}[x, y]$ such that $\mathbf{V}(f) = \gamma$. Then $p \in \gamma$ is a singular point of γ if and only if $\nabla f(p) = 0$ (that is, if the gradient is a vector of two zeros). We denote the set of singular points of γ as γ_{sing}.

To see why we require f to be of a minimum degree, let $\gamma \subset \mathbb{R}^2$ be the x-axis. It can be easily verified that γ has no singular points by writing $\gamma = \mathbf{V}(y)$ (and this also fits the above intuitive definition). If instead we rely on $\gamma = \mathbf{V}(y^2)$, we get the gradient $\nabla y^2 = (0, 2y)$ which is 0 at every point of γ. Thus, by using y^2 instead of y we get the false impression that every point of γ is singular.

A polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ is said to be *square-free* if in the factorization of f into irreducible factors, no factor has a multiplicity larger than one. Let $f \in \mathbb{R}[x_1, \ldots, x_d]$ be a square-free polynomial and let g be an irreducible factor of f that depends on x_i. Then $\frac{\partial f}{\partial x_i}$ is not divisible by g. Indeed, write $f = g \cdot h$ for some $h \in \mathbb{R}[x_1, \ldots, x_d]$ that does not have g as a factor, and notice that

$$\frac{\partial f}{\partial x_i} = \frac{\partial g}{\partial x_i} \cdot h + g \cdot \frac{\partial h}{\partial x_i}.$$

Since the second summand is divisible by has g but the first summand is not, this expression does not have g as a factor.

We now rely on square-free polynomials to establish that a curve cannot have too many singular points.

Theorem 2.6. *Let $\gamma \subset \mathbb{R}^2$ be an irreducible curve of degree k. Then γ_{sing} is a set of at most $k(k-1)$ points.*

Proof. Consider a minimum-degree polynomial $f \in \mathbb{R}[x, y]$ such that $\mathbf{V}(f) = \gamma$. Since removing repeated factors from f does not affect $\mathbf{V}(f)$, we know that f is square-free. Since γ is irreducible, we have that f is irreducible. Without loss of generality, we assume that f contains the variable x.

By the definition of a singular point, both $f_x = \frac{\partial f}{\partial x}$ and f vanish on every singular point of γ. Since f is square-free, it has no common components with f_x. By Bézout's theorem (Theorem 2.5), $\mathbf{V}(f) \cap \mathbf{V}(f_x)$ consists of at most $k(k-1)$ points. Thus, γ has at most $k(k-1)$ singular points. $\square$

As already mentioned, the above intuitive definition of a singular point is not completely accurate. While the three cases in that definition always lead to singular points, there are singular points that do not fit any of these cases. As an example, consider the polynomial $f = y^3 + 2x^2y - x^4 \in \mathbb{R}[x, y]$. The curve $\gamma = \mathbf{V}(f)$ is depicted in Figure 2.3. Notice that γ is an irreducible variety that does not intersect itself and has a well-defined tangent line at every point. Nonetheless, it can be easily verified that $\nabla f(0, 0) = (0, 0)$ so the origin is a singular point of γ. (For a discussion of this phenomenon, see for example [12, Section 3.3].)
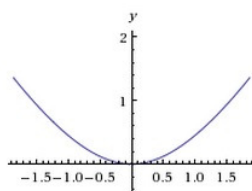
Figure 2.3: The variety $\mathbf{V}(y^3 + 2x^2y - x^4) \subset \mathbb{R}^2$.

Hopefully the definition of a *connected component* of a variety is sufficiently intuitive not to require a rigorous definition. We conclude this chapter by bounding the number of connected components that a variety in $\mathbb{R}^2$ can have.

Theorem 2.7 (Harnack's curve theorem). *Let $f \in \mathbb{R}[x, y]$ be a polynomial of degree k. Then the number of connected components of $\mathbf{V}(f)$ is $O(k^2)$.*

The exact bound of Harnack's theorem is $1 + \binom{k-1}{2}$. The proof that is presented here yields a slightly worse bound.

Proof. We may assume that f is square-free, since removing repeated factors does not change $\mathbf{V}(f)$. Every bounded connected component of $\mathbf{V}(f)$ has at least two extreme points in the x-direction (its leftmost and rightmost points). Such a point $p \in \mathbf{V}(f)$ satisfies $f(p) = \frac{\partial f}{\partial y}(p) = 0$. Since f is square-free, it has no common components with $f_y = \frac{\partial f}{\partial y}$. Thus, by Bézout's theorem (Theorem 2.5) $\mathbf{V}(f) \cap \mathbf{V}(f_y)$ has at most $k(k-1)$ points. This in turn implies that the number of bounded connected components of $\mathbf{V}(f)$ is at most $k(k-1)/2$.

To bound the number of unbounded connected components of $\mathbf{V}(f)$, we consider a sufficiently large constant c so that only the unbounded connected components intersect the lines $\mathbf{V}(x - c)$, $\mathbf{V}(x + c)$, $\mathbf{V}(y - c)$, and $\mathbf{V}(y + c)$. By Bézout's theorem, $\mathbf{V}(f)$ intersects each of those lines in at most k points, so there are at most $4k$ unbounded connected components. $\qquad\square$

2.3   Exercises

Problem 2.1. For each of the following sets, prove that it is not a variety:
(a) The sine wave $\{(x, \sin x) : x \in \mathbb{R}\}$ (hint: consider lines that intersect this set).
(b) The disc $\{(a, b) \in \mathbb{R}^2 : \sqrt{a^2 + b^2} \leq 1\}$.
(c) Every point set in $\mathbb{R}^2$ whose cardinality is countably infinite.

Problem 2.2. Let $\gamma \subset \mathbb{R}^2$ be a curve of degree k. Prove that the number of singular points of γ is $O(k^2)$ (Theorem 4.8 applies only to the case where γ is irreducible).

Problem 2.3. Prove or disprove:
(a) If $U, V \subset \mathbb{R}^2$ are varieties then the cartesian product $U \times V \subset \mathbb{R}^4$ is a variety.
(b) If $U \subset \mathbb{R}^3$ is a variety then the projection of U onto the xy-plane is a variety (that is, the set $\{(x, y) \in \mathbb{R}^2 \ : \ (x, y, z) \in U \text{ for some } z \in \mathbb{R}\}$).
(c) The complex variant of Theorem 2.5 (where $f, g \in \mathbb{C}[x, y]$) immediately implies the real variant.

Problem 2.4. Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^2$ and let $f \in \mathbb{R}[x, y]$ satisfy $\mathbf{V}(f) = \mathcal{P}$. Prove that $\deg f = \Omega(n^{1/2})$.

Chapter 3

Polynomial Partitioning

In this chapter we finally start to discuss the polynomial method. We introduce the polynomial partitioning theorem, use this theorem to derive a bound for incidences with general algebraic curves in $\mathbb{R}^2$, and then prove the theorem. Although we only consider incidences in $\mathbb{R}^2$, we prove the polynomial partitioning theorem in $\mathbb{R}^d$. This simple proof is identical in any dimension and does not require any of the properties of varieties in $\mathbb{R}^d$ that are presented in the following chapter.

3.1 The polynomial partitioning theorem

Consider a set $\mathcal{P}$ of m points in $\mathbb{R}^d$. For any $r > 1$, we say that $f \in \mathbb{R}[x_1, \ldots, x_d]$ is an *r-partitioning polynomial* for $\mathcal{P}$ if every connected component of $\mathbb{R}^d \setminus \mathbf{V}(f)$ contains at most m/r^d points of $\mathcal{P}$.[1] Notice that there is no restriction on the number of points of $\mathcal{P}$ that lie in $\mathbf{V}(f)$. Figure 3.1 depicts a 2-partitioning polynomial for a set of 12 points in $\mathbb{R}^2$.

The following result is due to Guth and Katz [51].

Theorem 3.1 (Polynomial partitioning [51]). *Let $\mathcal{P}$ be a set of m points in $\mathbb{R}^d$. Then for every $1 < r \leq m$, there exists an r-partitioning polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ of degree $O(r)$.*

To estimate the number of cells in such a partition, we rely on the following theorem.

---

[1]Currently there is no standard definition for an r-partitioning polynomial. Some authors define it to be a polynomial with every cell of $\mathbb{R}^d \setminus \mathbf{V}(f)$ containing at most m/r points, others use the notation $1/r$-partitioning polynomial, and so on. We chose the definition that in our opinion is the easiest one to work with.
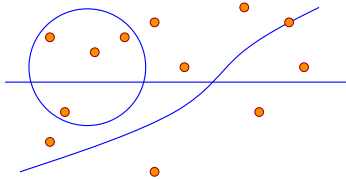
Figure 3.1: A 2-partitioning polynomial for a set of 12 points in $\mathbb{R}^2$.

Theorem 3.2 (Warren [104]). *For a polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ of degree k, the number of connected components of $\mathbb{R}^d \setminus \mathbf{V}(f)$ is $O_d\left(k^d\right)$.*

By Theorem 3.2, a polynomial of degree $O(r)$ yields $O(r^d)$ cells, which is asymptotically the minimum required to have at most m/r^d points in each cell. Thus, Theorem 3.1 is asymptotically tight (usually most of the points will indeed be in the cells and not on the partition itself).

In some degenerate cases, it might happen that all of the cells of the partition are empty. For example, consider a set $\mathcal{P}$ of m points in $\mathbb{R}^2$, all on the x-axis, and let f be an r-partitioning polynomial $\mathcal{P}$ (for r asymptotically smaller than m). If $\mathbf{V}(f)$ does not contain the x-axis, then by Bézout's theorem $\mathbf{V}(f)$ intersects the x-axis in $O(r)$ points. This in turn implies that most of the points of $\mathcal{P}$ are in $O(r)$ cells, contradicting the property that every cell contains at most m/r^2 points of $\mathcal{P}$. Therefore $\mathbf{V}(f)$ contains the x-axis and all of the cells are empty.

3.2 Incidences with algebraic curves in $\mathbb{R}^2$

Consider a point set $\mathcal{P}$ and a set of curves Γ, both in $\mathbb{R}^2$. The *incidence graph* of $\mathcal{P} \times \Gamma$ is a bipartite graph $G = (V_1 \cup V_2, E)$, where the vertices of V_1 correspond to the points of $\mathcal{P}$, the vertices of V_2 correspond to the curves of Γ, and an edge $(v_i, v_j) \in E$ implies that the point that corresponds to v_i is incident to the curve that corresponds to v_j; that is, E can be thought as the set of incidences in $\mathcal{P} \times \Gamma$. An example is depicted in Figure 3.2.

Recall that $K_{s,t}$ is a complete bipartite graph with s vertices on one side and t vertices on the other. Our goal in this section is to prove the following theorem (variants of this result originally appeared in [22, 70]).

Theorem 3.3. *Let $\mathcal{P}$ be a set of m points and let Γ be a set of n distinct irreducible algebraic curves of degree at most k, both in $\mathbb{R}^2$. If the incidence graph of $\mathcal{P} \times \Gamma$ contains no copy of $K_{s,t}$, then*

$$I(\mathcal{P}, \Gamma) = O_{s,t,k}\left(m^{\frac{s}{2s-1}} n^{\frac{2s-2}{2s-1}} + m + n\right).$$
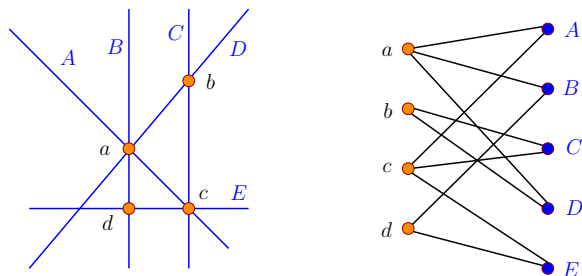
Figure 3.2: A point-line configuration and its incidence graph.

It is straightforward to generalize Theorem 3.3 to sets of curves that are neither distinct nor irreducible. We include these restrictions only to simplify the analysis (see also Problem 3.5). To emphasize the strength of Theorem 3.3, we consider some common types of curves:

- If Γ is a set of lines, since two lines intersect in at most one point, the incidence graph contains no copy of $K_{2,2}$. That is, Theorem 3.3 generalizes the Szemerédi-Trotter theorem.

- If Γ is a set of unit circles, the incidence graph contains no copy of $K_{2,3}$. That is, Theorem 3.3 generalizes the best bound for the unit distances problem (Theorem 1.6).

- If Γ is a set of arbitrary circles, the incidence graph contains no copy of $K_{3,2}$. In this case we obtain the bound $I(\mathcal{P}, \Gamma) = O\left(m^{3/5}n^{4/5} + m + n\right)$.

It is known that Theorem 3.3 is not tight in various cases, such as for parabolas and arbitrary circles. For some of the current best bounds and common conjectures, see Section 3.6.

We begin by proving a weaker incidence bound, which is purely combinatorial. This bound can be seen as a special case of the Kővari–Sós–Túran theorem from Extremal Graph Theory (e.g., see [64, Section 4.5]).

Lemma 3.4. *Let $\mathcal{P}$ be a set of m points and let Γ be a set of n curves, both in $\mathbb{R}^2$. If the incidence graph of $\mathcal{P} \times \Gamma$ contains no copy of $K_{s,t}$, then*

$$I(\mathcal{P}, \Gamma) = O_{s,t}\left(mn^{1-\frac{1}{s}} + n\right).$$

Proof. Let T be the set of $(s+1)$-tuples $(a_1, \ldots, a_s, \gamma)$ such that $a_1, \ldots, a_s \in \mathcal{P}$, $\gamma \in \Gamma$, and $a_1, \ldots, a_s \in \gamma$. We prove the lemma by double counting $|T|$.

On one hand, there are $\binom{m}{s}$ subsets of s points of $\mathcal{P}$, and every such subset is contained in at most $t - 1$ curves of Γ. That is,

$$|T| \le \binom{m}{s}(t - 1) = O_{s,t}\left(m^s\right). \tag{3.1}$$

Let $\Gamma = \{\gamma_1, \cdots, \gamma_n\}$. For each $\gamma_i \in \Gamma$ put $d_i = |\mathcal{P} \cap \gamma_i|$, so that $I(\mathcal{P}, \Gamma) = \sum_{i=1}^n d_i$. We have

$$|T| = \sum_{i=1}^n \binom{d_i}{s} = \Omega_s\left(\sum_{i=1}^n (d_i - s)^s\right).$$

By applying Hölder's inequality (see the "Notation and inequalities" part of the introduction) with $a_i = d_i - s$ and $b_i = 1$ for every $1 \le i \le n$, and with $p = s$, we get

$$\sum_{i=1}^n (d_i - s) \le \left(\sum_{i=1}^n (d_i - s)^s\right)^{1/s} \left(\sum_{i=1}^n 1\right)^{(s-1)/s} = \left(\sum_{i=1}^n (d_i - s)^s\right)^{1/s} n^{(s-1)/s}.$$

Since $I(\mathcal{P}, \Gamma) = \sum_{i=1}^n d_i$, we get

$$|T| = \Omega\left(\sum_{i=1}^n (d_i - s)^s\right) = \Omega\left(\frac{(I(\mathcal{P}, \Gamma) - sn)^s}{n^{s-1}}\right). \tag{3.2}$$

By combining (3.1) and (3.2), we obtain

$$\frac{(I(\mathcal{P}, \Gamma) - sn)^s}{n^{s-1}} = O_{s,t}\left(m^s\right).$$

Hence $I(\mathcal{P}, \Gamma) = O_{s,t}\left(mn^{(s-1)/s} + n\right)$, as asserted. □

To prove Theorem 3.3, we partition $\mathbb{R}^2$ into cells using Theorem 3.1, and then applying the bound of Lemma 3.4 separately in each cell. That is, we amplify the weak bound of Lemma 3.4 by combining it with polynomial partitioning.

We first present some intuition for why the above approach works. One way to think of the bound of Lemma 3.4 is that on average each point of $\mathcal{P}$ contributes $O\left(n^{(s-1)/s}\right)$ incidences (where n is the number of curves). On the other hand, when applying Lemma 3.4 separately in each cell, it is as if every point p contributes $O\left(n_p^{(s-1)/s}\right)$ incidences, where n_p is the number of curves that intersect the specific cell that contains p. Since a curve cannot intersect many cells, we expect $\sum_p n_p$ to be significantly smaller than $mn^{(s-1)/s}$.

Proof of Theorem 3.3. By Theorem 3.1, there exists an r-partitioning polynomial $f \in \mathbb{R}[x, y]$ for $\mathcal{P}$ of degree $O(r)$. We may assume that f is a minimum-degree polynomial that defines $\mathbf{V}(f)$. In particular, this means that f is square-free. The value of r will be determined below.

Let c denote the number of cells in (i.e., connected components of) $\mathbb{R}^2 \setminus \mathbf{V}(f)$. We denote by $\mathcal{P}_0 = \mathbf{V}(f) \cap \mathcal{P}$ the set of points of $\mathcal{P}$ that are contained in $\mathbf{V}(f)$. Similarly, we denote by Γ_0 the set of curves of Γ that are fully contained in $\mathbf{V}(f)$. For $1 \leq i \leq c$, let $\mathcal{P}_i$ denote the set of points that are contained in the i-th cell and let Γ_i denote the set of curves of Γ that intersect the i-th cell. Notice that

$$I(\mathcal{P}, \Gamma) = I(\mathcal{P}_0, \Gamma_0) + I(\mathcal{P}_0, \Gamma \setminus \Gamma_0) + \sum_{i=1}^{c} I(\mathcal{P}_i, \Gamma_i).$$

We bound each of these three expressions separately.

We begin with $\sum_{i=1}^{c} I(\mathcal{P}_i, \Gamma_i)$, for which we require an upper bound on the number of cells c. By Theorem 3.2, we have $c = O(r^2)$. We set $m_i = |\mathcal{P}_i| \leq m/r^2$ and $n_i = |\Gamma_i|$. By applying Lemma 3.4 separately in each cell, we have

$$\sum_{i=1}^{c} I(\mathcal{P}_i, \Gamma_i) = O_{s,t}\left(\sum_{i=1}^{c}(m_i n_i^{\frac{s-1}{s}} + n_i)\right) = O_{s,t}\left(\frac{m}{r^2}\sum_{i=1}^{c} n_i^{\frac{s-1}{s}} + \sum_{i=1}^{c} n_i\right).$$

We claim that any curve $\gamma \in \Gamma$ intersects $O_k(r)$ cells of the partition. When traveling across a connected component of γ, to enter a new cell of the partition we must first intersect $\mathbf{V}(f)$. By Bézout's theorem, the number of intersection points between a curve $\gamma \in \Gamma$ and $\mathbf{V}(f)$ is $O_k(r)$. By Harnack's curve theorem (Theorem 2.7), γ has $O_k(1)$ connected components. These two bounds do not suffice to claim that γ intersects $O_k(r)$ cells, since in any intersection point between γ and $\mathbf{V}(f)$, the curve γ may split into several cells (e.g., See Figure 3.3). Consider such an intersection point p and let C_p be a circle that is centered at p and of a sufficiently small radius. By Bézout's theorem, γ and C_p intersect in at most $2k$ points. Thus, γ may split into $O_k(1)$ cells in an intersection point with the partition. This completes our claim that γ intersects $O_k(r)$ cells of the partition.

Since any curve $\gamma \in \Gamma$ intersects $O_k(r)$ cells of the partition, we have $\sum_{i=1}^{c} n_i = O_k(nr)$. By Hölder's inequality, we have

$$\sum_{i=1}^{c} n_i^{\frac{s-1}{s}} \leq \left(\sum_{i=1}^{c} n_i\right)^{\frac{s-1}{s}}\left(\sum_{i=1}^{c} 1\right)^{\frac{1}{s}} = O_k\left((nr)^{\frac{s-1}{s}}\left(r^2\right)^{\frac{1}{s}}\right) = O_k\left(n^{\frac{s-1}{s}} r^{\frac{s+1}{s}}\right).$$
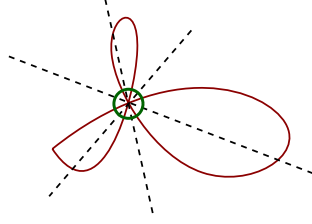
Figure 3.3: The dashed lines represent the partition. In an intersection with the partition, the red curve splits into six cells.

Combining the above implies

$$\sum_{i=1}^{c} I(\mathcal{P}_i, \Gamma_i) = O_{s,t,k} \left(\frac{m}{r^2} \sum_{i=1}^{c} n_i^{\frac{s-1}{s}} + \sum_{i=1}^{c} n_i \right) = O_{s,t,k} \left(\frac{mn^{\frac{s-1}{s}}}{r^{\frac{s-1}{s}}} + nr \right). \qquad (3.3)$$

Next, consider a curve $\gamma \in \Gamma \setminus \Gamma_0$. Since the number of intersection points between γ and $\mathbf{V}(f)$ is $O_k(r)$, we get

$$I(\mathcal{P}_0, \Gamma \setminus \Gamma_0) = O_k(nr). \qquad (3.4)$$

It remains to bound $I(\mathcal{P}_0, \Gamma_0)$. Notice that $\mathbf{V}(f)$ consists of $O(r)$ one-dimensional (irreducible) components. Since the curves of Γ are irreducible and distinct, each component of $\mathbf{V}(f)$ corresponds to at most one curve of Γ. Recall that a point that is contained in more than one component of $\mathbf{V}(f)$ is a singular point of $\mathbf{V}(f)$. Thus, every regular point of $\mathbf{V}(f)$ is incident to at most one curve of Γ_0. That is, there are $O(m)$ incidences between curves of Γ_0 and points of $\mathcal{P}_0$ that are regular points of $\mathbf{V}(f)$.

Since it is impossible for both first partial derivatives of f to be identically zero, without loss of generality we assume that $f_x = \frac{\partial f}{\partial x}$ is not identically zero. By definition, f_x vanishes on every singular point of $\mathbf{V}(f)$. Since f is square-free, it has no common components with f_x. Consider $\gamma \in \Gamma_0$, and notice that γ and $\mathbf{V}(f_x)$ also have no common components. By Bézout's theorem, $\gamma \cap \mathbf{V}(f_x)$ consists of $O_k(r)$ points. That is, γ is incident to $O_k(r)$ singular points of $\mathbf{V}(f)$. By summing the above over every $\gamma \in \Gamma_0$, we have

$$I(\mathcal{P}_0, \Gamma_0) = O_k(nr + m). \qquad (3.5)$$

By combining (3.3), (3.4), and (3.5), we obtain

$$I(\mathcal{P}, \Gamma) = O_{s,t,k} \left(\frac{mn^{\frac{s-1}{s}}}{r^{\frac{s-1}{s}}} + nr + m \right).$$

It remains to find the value of r that minimizes the above bound. Since the first term in this bound is decreasing in r while the second is increasing in r, the optimal bound is obtained when both terms are equivalent. Thus, the optimal value for r is $\Theta\left(m^{\frac{s}{2s-1}}/n^{\frac{1}{2s-1}}\right)$. Setting this value immediately implies the assertion of the theorem.

One minor issue: When $m = O(n^{1/s})$ we might have $m^{\frac{s}{2s-1}}/n^{\frac{1}{2s-1}} < 1$, which may lead to an invalid value of r. Fortunately, in this case Lemma 3.4 implies the bound $I(\mathcal{P},\Gamma) = O_{s,t}(n)$. $\hfill\square$

We briefly repeat the main steps for deriving an incidence bound using polynomial partitioning, since several variants of this approach appear in the following chapters.

- First, we obtain a "weak" incidence bound by using a combinatorial argument (see Lemma 3.4).
- We partition the space into cells by using a partitioning polynomial.
- We apply the weak incidence bound separately in each cell of the partition.
- Finally, we bound the number of incidences on the partition itself.

3.3   Proving the polynomial partitioning theorem

In this section we prove the polynomial partitioning theorem. We first repeat the statement of this theorem.

Theorem 3.1. *Let $\mathcal{P}$ be a set of m points in $\mathbb{R}^d$. Then for every $1 < r \leq m$, there exists an r-partitioning polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ of degree $O(r)$.*

Intuitively, to prove the theorem we iteratively partition $\mathcal{P}$. That is, we first partition $\mathcal{P}$ into two disjoint sets $\mathcal{P}_1, \mathcal{P}_2 \subset \mathcal{P}$ that are not too large (and may not contain some points of $\mathcal{P}$). We then partition both $\mathcal{P}_1$ and $\mathcal{P}_2$ to obtain four sets $\mathcal{P}'_1, \mathcal{P}'_2, \mathcal{P}'_3, \mathcal{P}'_4$, etc. An example of this process is depicted in Figure 3.4.

$$\mathbf{V}(f_1) \qquad\qquad\qquad \mathbf{V}(f_2) \qquad\qquad\qquad \mathbf{V}(f_3)$$
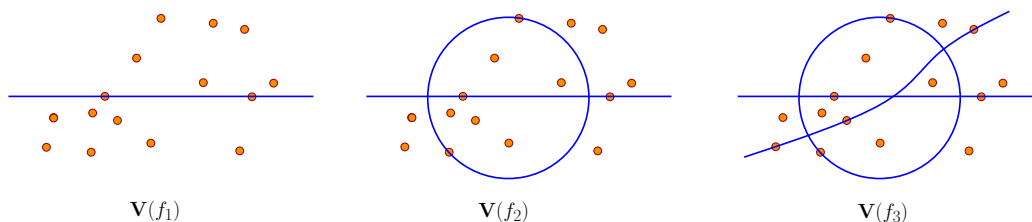
Figure 3.4: Repeatedly partitioning a set of 15 points in the plane. At the end of this process no cell contains more than two points.

A *hyperplane* in $\mathbb{R}^d$ is a variety that is defined by a linear equation (and thus looks like a copy of $\mathbb{R}^{d-1}$). A hyperplane h in $\mathbb{R}^d$ *bisects* a finite point set $\mathcal{P} \subset \mathbb{R}^d$ if each of the two open halfspaces bounded by h contains at most $|\mathcal{P}|/2$ points of $\mathcal{P}$. The bisecting hyperplane may contain any number of points of $\mathcal{P}$. The following is a discrete version of the *ham sandwich theorem* (e.g., see [63]).

Theorem 3.5. *Every d finite point sets $\mathcal{P}_1, \ldots, \mathcal{P}_d \subset \mathbb{R}^d$ can be simultaneously bisected by a hyperplane.*

A planar example of Theorem 3.5 is depicted in Figure 3.5. To iteratively partition $\mathcal{P}$, we can apply Theorem 3.5. However, after about $\log_2 d$ steps we obtain more than d sets of points, and can no longer apply the theorem. Indeed, it is not hard to find $d + 1$ sets of points that cannot be simultaneously bisected by a hyperplane. To overcome this difficulty, we instead use a discrete version of the *polynomial ham sandwich theorem*.
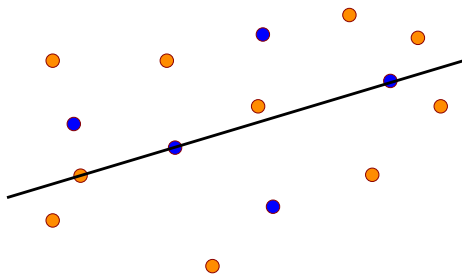
Figure 3.5: The line simultaneously bisects the set of blue points and the set of orange points.

A polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ bisects a finite point set $\mathcal{P} \subset \mathbb{R}^d$ if $f(p) > 0$ for at most $|\mathcal{P}|/2$ points $p \in \mathcal{P}$ and $f(p) < 0$ for at most $|\mathcal{P}|/2$ points $p \in \mathcal{P}$. The variety $\mathbf{V}(f)$ may contain any number of points of $\mathcal{P}$.

Theorem 3.6 (Stone and Tukey [95]). *Let $\mathcal{P}_1, \ldots, \mathcal{P}_t \subset \mathbb{R}^d$ be t finite point sets, and let D be an integer such that $\binom{D+d}{d} - 1 \geq t$. Then there exists a nonzero polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ of degree at most D that simultaneously bisects all of the sets $\mathcal{P}_i$.*

Proof. The number of monomials that a polynomial of degree at most D in $\mathbb{R}[x_1, \ldots, x_d]$ can have is $\binom{D+d}{d}$; this is illustrated in Figure 3.6. Let

$$U_D = \{(u_1, \ldots, u_d) \in \mathbb{Z}^d \ : \ 1 \leq u_1 + \cdots + u_d \leq D \ \text{ and } \ u_i \geq 0 \ \text{ for every } \ 1 \leq i \leq d\}.$$

Intuitively, this is the set of exponents of nonconstant monomials of degree at most D in $\mathbb{R}[x_1, \ldots, x_d]$. Set $m = |U_D|$ and note that $m = \binom{D+d}{d} - 1$. The *Veronese map* $\nu_D : \mathbb{R}^d \to \mathbb{R}^m$ is defined as

$$\nu_D(x_1, \ldots, x_d) := (x_1^{u_1} x_2^{u_2} \cdots x_d^{u_d})_{u \in U_D}.$$
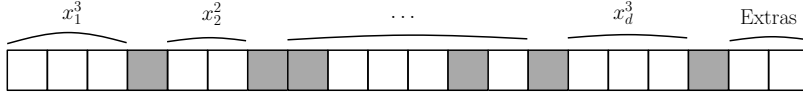
Figure 3.6: Every monomial of degree at most D in $x_1, \ldots, x_d$ corresponds to a unique choice of the d gray blocks out of a total of $D + d$ blocks.

Every coordinate in $\mathbb{R}^m$ corresponds to a nonconstant monomial of degree at most D in $\mathbb{R}[x_1, \ldots, x_d]$, and $\nu_D(\cdot)$ maps a point p in $\mathbb{R}^d$ to the tuple of the values of these monomials at p. For example, the Veronese map $\nu_2 : \mathbb{R}^2 \to \mathbb{R}^5$ is

$$\nu_2(x, y) = (x^2, xy, y^2, x, y).$$

For every $1 \le i \le t$, we set $\mathcal{P}'_i = \nu_D(\mathcal{P}_i)$. That is, every $\mathcal{P}'_i$ is a finite point set in $\mathbb{R}^m$. By the assumption on D, we have $m \ge t$. Thus, by Theorem 3.5 there exists a hyperplane $\Pi \subset \mathbb{R}^m$ that simultaneously bisects all of the sets $\mathcal{P}'_i$. We denote the coordinates of $\mathbb{R}^m$ as y_u (for each $u \in U$), so that Π can be defined by a linear equation of the form $h_0 + \sum_{u \in U} y_u h_u$, for a suitable set of constants $h_u \in \mathbb{R}$.

Returning to $\mathbb{R}^d$, we consider the polynomial $f(x_1, \ldots, x_d) = h_0 + \sum_{u \in U} h_u x_1^{u_1} x_2^{u_2} \cdots x_d^{u_d}$. For any point $a \in \mathbb{R}^d$ and $a' = \nu_D(a)$, we have that $h_0 + (h_u)_{u \in U} \cdot a' = f(a)$. That is, for every point $a \in \mathbb{R}^d$, $f(a) > 0$ (resp., $f(a) < 0$) if and only if $\nu_D(a)$ is in the positive side of Π (resp., in the negative side of Π). Since Π bisects every $\mathcal{P}'_i$, the polynomial f bisects every A_i. This concludes the proof since f is of degree at most D. $\square$

Guth and Katz relied on the polynomial ham sandwich theorem to derive the polynomial partitioning theorem.

Proof of Theorem 3.1. Theorem 3.6 implies the existence of a bisecting polynomial of degree at most $c_d t^{1/d}$, for a constant c_d depending only on d. To prove the theorem, we show that there exists a sequence of polynomials $f_1, f_2, \ldots$ such that the degree of f_j is smaller than $c_d 2^{(j+1)/d}/(2^{1/d} - 1)$ and every connected component of $\mathbb{R}^d \setminus \mathbf{V}(f_j)$ contains at most $m/2^j$ points of $\mathcal{P}$. An example is depicted in Figure 3.4. This would complete the proof since we can then choose $f = f_s$, where s is the minimum integer satisfying $2^s \ge r^d$.

We prove the existence of f_j by induction on j. The existence of f_1 is immediate from Theorem 3.5, so we move to the induction step. By the induction hypothesis, there exists a polynomial f_j of degree smaller than $c_d 2^{(j+1)/d}/(2^{1/d} - 1)$ such that every connected component of $\mathbb{R}^d \setminus \mathbf{V}(f_j)$ contains at most $m/2^j$ points of $\mathcal{P}$. Since $|\mathcal{P}| = m$, the number t of connected components of $\mathbb{R}^d \setminus \mathbf{V}(f_j)$ that contain more than $m/2^{j+1}$ points of $\mathcal{P}$ is smaller than 2^{j+1}. Let $\mathcal{P}_1, \ldots, \mathcal{P}_r \subset \mathcal{P}$ be the subsets of $\mathcal{P}$ that are contained in each of these connected components (that is, $|\mathcal{P}_1|, \ldots, |\mathcal{P}_r| > m/2^{j+1}$). By Theorem 3.6, there is a polynomial g_j of degree smaller than $c_d 2^{(j+1)/d}$ that simultaneously bisects every $\mathcal{P}_j$. We can set $f_{j+1} = f_j \cdot g_j$, since every connected component of $\mathbb{R}^d \setminus \mathbf{V}(f_j \cdot g_j)$ contains at most $m/2^{j+1}$ points of $\mathcal{P}$ and $f_j \cdot g_j$ is a polynomial of degree smaller than

$$\frac{c_d 2^{(j+1)/d}}{2^{1/d} - 1} + c_d 2^{(j+1)/d} = c_d 2^{(j+1)/d} \cdot \left(\frac{1}{2^{1/d} - 1} + 1 \right) = \frac{c_d 2^{(j+2)/d}}{2^{1/d} - 1}.$$

This completes the induction step, and thus also the proof. $\qquad\square$

3.4 Curves containing lattice points

We conclude this chapter with a simple application of Theorem 3.3. Let $\mathcal{G}$ be a $\sqrt{n} \times \sqrt{n}$ section of the integer lattice in $\mathbb{R}^2$. It is easy to show that any constant-degree algebraic curve is incident to $O(\sqrt{n})$ points of $\mathcal{G}$ (e.g., by using Bézout's theorem). This bound is tight, since a line can pass through $\Theta(\sqrt{n})$ points of $\mathcal{G}$. We now show that every non-line constant-degree algebraic curve passes through an asymptotically smaller number of lattice points (it is based on ideas similar to those of Iosevich [56]). In Problem 3.13 we derive the improved bound $O_k(n^{1/3})$, while the stronger bound $O_k(n^{1/(2k)})$ was derived by Bombieri and Pila [13], relying on number-theoretic methods.

Claim 3.7. *Let $\mathcal{G}$ be a $\sqrt{n} \times \sqrt{n}$ section of the integer lattice in $\mathbb{R}^2$, and let γ be an irreducible algebraic curve of degree $k \geq 2$. Then γ contains $O_k \left(n^{k^2/(2k^2+1)} \right)$ points of $\mathcal{G}$.*

Proof. Let x denote the number of points of $\mathcal{G}$ that are incident to γ, let p be a point of $\mathcal{G}$ that is incident to γ, and let $\mathbb{T}$ denote the set of translations of $\mathbb{R}^2$ that take p to another point of $\mathcal{G}$. We apply each of the translations of $\mathbb{T}$ on γ to obtain a set Γ of n copies of γ. An example is depicted in Figure 3.7(a,b).

Some of the translated copies of γ might contain fewer than x points of $\mathcal{G}$. To fix this, we also apply each translation of $\mathbb{T}$ on the points of $\mathcal{G}$ (see Figure 3.7(c)).
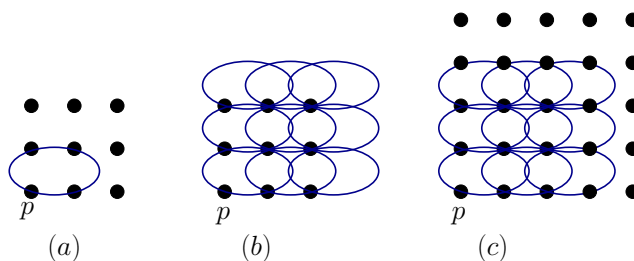
Figure 3.7: (a) A curve γ containing lattice points. (b) Applying the translations of $\mathbb{T}$ on γ. (c) Applying the translations of $\mathbb{T}$ on $\mathcal{G}$.

Notice that this results in a set $\mathcal{G}'$ of less than $4n$ distinct lattice points. To complete the proof, we double count $I(\mathcal{G}', \Gamma)$. After inserting the additional points, each of the n copies of γ contains at least x points of $\mathcal{G}$. That is, $I(\mathcal{G}', \Gamma) = \Omega(nx)$.

Notice that two translated copies of an irreducible curve cannot have a common component. Thus, by Bézout's theorem any two curves of Γ have at most k^2 points in common. By Theorem 3.3 with $s = k^2 + 1$, we obtain the bound $I(\mathcal{G}', \Gamma) = O_k(n^{(3k^2+1)/(2k^2+1)})$. Combining our two bounds for $I(\mathcal{G}', \Gamma)$ immediately implies the assertion of the claim. $\qquad\square$

3.5   Exercises

Problem 3.1. Find a set $\mathcal{P}$ of n points in $\mathbb{R}^2$, a point $q \notin \mathcal{P}$, and an integer r, such that there is no r-partitioning polynomial of $\mathcal{P}$ that does not contain q. That is, show that Theorem 3.1 is no longer true when also asking $\mathbf{V}(f)$ not to contain a specific point.

Problem 3.2. Show that Theorem 3.1 is no longer true when asking f to be irreducible. That is, find a set $\mathcal{P}$ of n points in $\mathbb{R}^2$ and an integer r such that there is no irreducible r-partitioning polynomial of $\mathcal{P}$.

Problem 3.3. Let $\mathcal{P}$ be a set of m points and let $\mathcal{C}$ be a set of n circles, both in $\mathbb{R}^2$. Theorem 3.3 implies $I(\mathcal{P}, \Gamma) = O\left(m^{3/5}n^{4/5} + m + n\right)$.
(a) Derive a stronger bound for the case where the centers of all the circles are on the x-axis.
(b) Derive a stronger bound for the case where no line contains more than 1,000 circle centers.

Problem 3.4. For $m \leq n$, let $\mathcal{P}$ be a set of m points on the x-axis and let $\mathcal{P}'$ be a

set of n points. Let $D(\mathcal{P}, \mathcal{P}')$ denote the number of distinct distances between $\mathcal{P}$ and $\mathcal{P}'$. That is, we only consider distances between pairs of points from $\mathcal{P} \times \mathcal{P}'$. Prove that $D(\mathcal{P}, \mathcal{P}') = \Omega(\sqrt{mn})$. Hint: Recall the proof of Claim 1.11 and use Problem 3.3.

Problem 3.5. Prove that Theorem 3.3 remains valid also after removing the restrictions about the curves being irreducible and distinct (hint: This can be done by treating the proof of Theorem 3.3 as a black box, rather than changing the proof of this theorem).

Problem 3.6. Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^2$, let $D \geq 2$, and let ν_D be the Veronese map of degree D from $\mathbb{R}^2$. Prove that $\nu_D(\mathcal{P}) = \{\nu_D(p) : p \in \mathcal{P}\}$ is in convex position. (Recall that a point set is in convex position if every $p \in \mathcal{P}$ can be separated from $\mathcal{P} \setminus \{p\}$ by a hyperplane.)

Problem 3.7. *Radon's theorem* states that any set of $d + 2$ points in $\mathbb{R}^d$ can be partitioned into two disjoint subsets whose convex hulls intersect. We say that a polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ *separates* two finite point sets $\mathcal{P}_1, \mathcal{P}_2 \subset \mathbb{R}^d$ if $f(p) > 0$ for every $p \in \mathcal{P}_1$ and $f(q) < 0$ for every $q \in \mathcal{P}_2$. It is known that the convex hulls of two finite point sets are disjoint if and only if no hyperplane separates them.

We are given a set $\mathcal{P}$ of n points in $\mathbb{R}^d$ such that n is much larger than d. Use Radon's theorem to prove that $\mathcal{P}$ can be partitioned into two disjoint subsets $\mathcal{P}_1, \mathcal{P}_2 \subset \mathcal{P}$ such that all of the polynomials that separate $\mathcal{P}_1$ and $\mathcal{P}_2$ are of degree $\Omega_d(n^{1/d})$ (you may rely on the statements in the first paragraph without proving them).

Problem 3.8. Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^2$. Prove that $\mathcal{P}$ spans $\Omega(n^{3/4})$ distinct distances. (Hint: Recall the proof of Claim 1.11.)

Problem 3.9. Prove that the maximum number of isosceles triangles that can be determined by a set of n points in $\mathbb{R}^2$ is $O\left(n^{7/3}\right)$. While it is possible to solve the problem by using point–line incidences, you are asked to solve it by reducing the problem to point–circle incidences.

Problem 3.10. Show that the polynomial partitioning theorem (Theorem 3.1) remains valid after adding the following restriction: No monomial of f contains a variable with an exponent larger than $r/2$.

Problem 3.11. Let $\mathcal{P}$ be a set of m points and let $\mathcal{L}$ be a set of n lines, both in $\mathbb{R}^2$. Prove that there exist $\mathcal{P}' \subset \mathcal{P}$ and $\mathcal{L}' \subset \mathcal{L}$ such that $|\mathcal{P}'| = \Theta(m)$, $|\mathcal{L}'| = \Theta(n)$, and $I(\mathcal{P}', \mathcal{L}') = 0$. (Hint: Consider a large constant r.)

Problem 3.12. Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^2$. Prove that there exists $\mathcal{P}' \subset \mathcal{P}$ such that $|\mathcal{P}'| = \Theta(n)$ and no unit distances are spanned by $\mathcal{P}'$. (Hint: How is this

related to Problem 3.11?)

Problem 3.13. Your goal in this problem is to improve the bound of Claim 3.7 to $O_k\left(n^{1/3}\right)$. The following paragraph suggests one approach for doing this.

Construct the sets $\mathcal{G}'$ and Γ as in the original proof, and then move to a dual plane, as follows. Recall that every curve $\gamma_j \in \Gamma$ is a translation of γ, which can be decomposed into a horizontal translation followed by a vertical translation. Instead of γ_j, consider a point v_j whose x-coordinate is the distance of the horizontal translation, and whose y-coordiante is the distance of the vertical one. Replace every point $p \in \mathcal{G}'$ with the set S_p of the points in $\mathbb{R}^2$ that correspond to translations of γ that are incident to p (we do not refer only to points that correspond to curves in Γ, but rather to any point that parameterizes a translated copy of γ incident to p). How does the set S_p look like?

Problem 3.14. The bound of Theorem 3.3 is of the form $O_{s,t,k}(\cdot)$. We wish to remove t from the subscript. That is, we wish to find the dependency of the bound in t (for example, this is useful when $t = \log m$ or $t = n^\alpha$ for some $0 < \alpha < 1$). Revise the proof of Theorem 3.3 accordingly. There is no need to rewrite the entire proof — focus on the parts that need to be revised.

Problem 3.15. Let $\mathcal{P}$ be a set of m points and let $\mathcal{L}$ be a set of n lines, both in $\mathbb{R}^d$. By Theorem 3.1, there exists a partition with $O(r^d)$ cells, each containing at most m/r^d points of $\mathcal{P}$ (for some $1 < r < m$). Unfortunately, for the application that we have in mind, we also require the property that no cell is intersected by many lines of $\mathcal{L}$.

Show that we can further partition the existing $O(r^d)$ cells so that (i) every new cell is intersected by at most n/r^{d-1} lines of $\mathcal{L}$, and (ii) the number of cells remains $O(r^d)$. Do this by partitioning each cell C into several abstract subcells (i.e., different subcells do not necessarily correspond to different geometric areas), where each subcell of C consists of the same set of points as C but only of a subset of the lines. As before, every point-line incidence is required to appear in exactly one subcell (unless it is on the original partitioning, in which case it is in none of the cells and subcells).

3.6   Open problems

Theorem 3.3 provides a general point–curve incidence bound in $\mathbb{R}^2$. In Chapter 1 we saw that this bound is tight for the case of lines. However, the theorem is suspected not to be tight for almost any other case. The following appears to be a common conjecture.

Conjecture 3.8. *Let $\mathcal{P}$ be a set of n points and let Γ be a set of n algebraic curves of degree at most k, both in $\mathbb{R}^2$. If the incidence graph of $\mathcal{P} \times \Gamma$ contains no copy of $K_{s,t}$, then*

$$I(\mathcal{P}, \Gamma) = O_{s,t,k}\left(n^{4/3}\right).$$

Conjecture 3.8 is known to be false when the number of curves is significantly larger than the number of points. For example, there exists a set of m points and a set of n parabolas, both in $\mathbb{R}^2$, with $\Theta\left(m^{1/2}n^{5/6}\right)$ incidences (see Problem 1.2). This expression is asymptotically larger than $m^{2/3}n^{2/3}$ when n is asymptotically larger than m.

As mentioned in Chapter 1, for any $\varepsilon > 0$ it is conjectured that the number of incidences between n points and n unit circles is $O(n^{1+\varepsilon})$. This might also be the case for several other variants, such as the case of the degenerate hyperbolas that were described in Problem 1.3. Currently no method is known for obtaining a bound asymptotically smaller than $O(n^{4/3})$.

When $s > 2$, the *lens cutting* method yields bounds that are somewhat stronger than the ones of Theorem 3.3. The following theorem contains the best known bounds for these cases.

Theorem 3.9 ([83]). *Let $\mathcal{P}$ be a set of m points and let Γ be a set of n irreducible algebraic curves of degree at most k in $\mathbb{R}^2$. Assume that we can parameterize these curves using s parameters. Then for every $\varepsilon > 0$ we have*

$$I(\mathcal{P}, \Gamma) = O_{k,s,\varepsilon}\left(m^{\frac{2s}{5s-4}+\varepsilon}n^{\frac{5s-6}{5s-4}} + m^{2/3}n^{2/3} + m + n\right).$$

Finally, Figure 3.8 is a good recap for everything that we have seen in the past couple of sections.
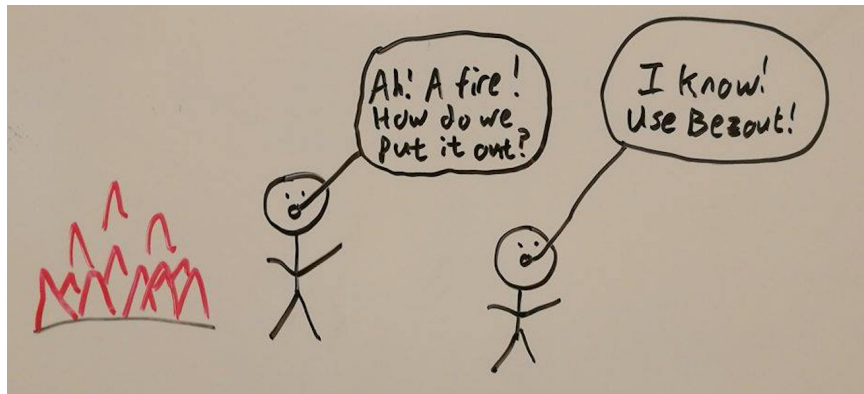
Figure 3.8: A drawing by Zachary Chase.

Chapter 4

Basic Real Algebraic Geometry in $\mathbb{R}^d$

> *"Every field has its taboos. In algebraic geometry the taboos are (1) writing a draft that can be followed by anyone but two or three of one's closest friends, (2) claiming that a result has applications, (3) mentioning the word "combinatorial", and (4) claiming that algebraic geometry existed before Grothendieck."* / Gian-Carlo Rota [79].

In this chapter we generalize to $\mathbb{R}^d$ several of the definitions and results that were studied in $\mathbb{R}^2$ in Chapter 2. For this, we first introduce the notion of a polynomial ideal and define the dimension of a variety in $\mathbb{R}^d$.

4.1   Ideals

In Chapter 2 we introduced varieties, which are the basic geometric objects of Algebraic Geometry. We now introduce polynomial ideals, which are the basic algebraic objects that are studied in this book. A subset $J \subseteq \mathbb{R}[x_1, \ldots, x_d]$ is an *ideal* if it satisfies:

- $0 \in J$.
- If $f, g \in J$ then $f + g \in J$.
- If $f \in J$ and $h \in \mathbb{R}[x_1, \ldots, x_d]$, then $f \cdot h \in J$.

As a first example of an ideal, consider a polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ and notice that the set $\{f \cdot h : h \in \mathbb{R}[x_1, \ldots, x_d]\}$ is an ideal. More generally, given $f_1, \ldots, f_k \in$

$\mathbb{R}[x_1, \ldots, x_d]$, the set

$$\langle f_1, \ldots, f_k \rangle = \left\{ \sum_{i=1}^{k} f_i \cdot h_i : h_1, \ldots, h_k \in \mathbb{R}[x_1, \ldots, x_d] \right\}$$

is an ideal. We say that this ideal is *generated* by $f_1, \ldots, f_k$. We also say that $\{f_1, \ldots, f_k\}$ is a *basis* of this ideal.[1]

We are specifically interested in ideals of varieties. Given a variety $U \subset \mathbb{R}^d$, the ideal of U is

$$\mathbf{I}(U) = \{f \in \mathbb{R}[x_1, \ldots, x_d] : f(a) = 0 \text{ for every } a \in U\}.$$

It can be easily verified that $\mathbf{I}(U)$ satisfies the three requirements in the definition of an ideal. This leads to the question: Given $f_1, \ldots, f_k \in \mathbb{R}[x_1, \ldots, x_d]$, is it always the case that $\langle f_1, \ldots, f_k \rangle = \mathbf{I}(\mathbf{V}(f_1, \ldots, f_k))$?

Claim 4.1. *Given $f_1, \ldots, f_k \in \mathbb{R}[x_1, \ldots, x_d]$, we have $\langle f_1, \ldots, f_k \rangle \subseteq \mathbf{I}(\mathbf{V}(f_1, \ldots, f_k))$ although equality need not occur.*

Proof. To see that the containment relation holds, we set $U = \mathbf{V}(f_1, \ldots, f_k)$. If $g \in \langle f_1, \ldots, f_k \rangle$ then by definition g vanishes on every point of U, and is thus in $\mathbf{I}(\mathbf{V}(f_1, \ldots, f_k))$.

To see that equality does not always hold, we set $f = x^2 + y^2$. We then have $\mathbf{V}(f) = \{(0,0)\} \subset \mathbb{R}^2$, $x \in \mathbf{I}(\mathbf{V}(f))$, and $x \notin \langle f \rangle$. $\square$

As we shall see, when defining a variety U it is often useful to use a basis of $\mathbf{I}(U)$ rather than an arbitrary set of polynomials that define U. We now inspect another connection between ideals and varieties.

Claim 4.2. *Let $U, W \subset \mathbb{R}^d$ be varieties. Then*
(a) $U \subsetneq W$ if and only if $\mathbf{I}(W) \subsetneq \mathbf{I}(U)$.
(b) $U = W$ if and only if $\mathbf{I}(W) = \mathbf{I}(U)$.

Proof. We only prove part (a); part (b) is proved in a similar manner.

First assume that $U \subset W$ and consider a polynomial $f \in \mathbf{I}(W)$. That is, f vanishes on every point of W. Since $U \subset W$ we get that f vanishes on every point of U, which implies $f \in \mathbf{I}(U)$. To see that this containment is proper, notice that there

---

[1]There are many different notations for an ideal generated by a set of polynomials. We use $\langle \cdot \rangle$ following the notation of [23].

must exist polynomials that vanish on U but not on W (otherwise we would have $U = W$).

Next, assume that $\mathbf{I}(W) \subset \mathbf{I}(U)$ and consider a point $p \in U$. Every polynomial of $\mathbf{I}(U)$ vanishes on p, which in turn implies that every polynomial of $\mathbf{I}(W)$ vanishes on p. Thus, we have $p \in W$. There must exist $q \in W \setminus U$, since otherwise we would have $\mathbf{I}(W) = \mathbf{I}(U)$. $\qquad \square$

We conclude this section with (a special case of) a classic result called the *ascending chain condition*.

Theorem 4.3. *Let $I_1 \subseteq I_2 \subseteq I_3 \subseteq \cdots$ be an infinite chain of ideals in $\mathbb{R}[x_1, \ldots, x_d]$ (or in a polynomial ring over any other field). Then there exists an integer $n \geq 1$ such that $I_n = I_{n+1} = I_{n+2} = \cdots$.*

4.2 Dimension

Consider an irreducible variety $U \subset \mathbb{R}^d$. One intuitive definition of the *dimension d'* of U, denoted $\dim U$, is the maximum integer such that there exists a sequence

$$U_0 \subset U_1 \subset \cdots \subset U_{d'} = U,$$

where all of the subsets are proper and all of the sets U_i are irreducible varieties. If $U \subset \mathbb{R}^d$ is a reducible variety with irreducible components $U_1, \ldots, U_k$, then we define $\dim U = \max_i \{\dim U_i\}$.

The above definition of dimension corresponds to our intuitive one. For example, any finite point set in $\mathbb{R}^d$ is a variety of dimension zero, any finite union of lines and circles in $\mathbb{R}^d$ is a variety of dimension one, and so on.

Claim 4.4. *Let U and V be two varieties in $\mathbb{R}^d$, both of dimension d' and with no common components. Then $\dim(U \cap V) < d'$.*

Proof. We first assume that U and V are both irreducible. Assume for contradiction that $\dim(U \cap V) \geq d'$. Let W be a component of $U \cap V$ of dimension at least d'. By definition, there exists a chain

$$W_0 \subset W_1 \subset \cdots \subset W_{d'-1} \subset W,$$

where all of the subsets are proper and all of the sets are irreducible varieties. Since U and V have no common components, W is a proper subset of U. Thus, the chain

$$W_0 \subset W_1 \subset \cdots \subset W_{d'-1} \subset W \subset U$$

also consists only of proper subsets and irreducible varieties. This contradicts the dimension of U being d', so $\dim(U \cap V) < d'$.

Next, we consider the case where U and V may be reducible. Let $U_1, \ldots, U_k$ be the components of U and let $V_1, \ldots, V_\ell$ be the components of V. Notice that $U \cap V = \bigcup_{\substack{1 \le i \le k \\ 1 \le j \le \ell}} (U_i \cap V_j)$. By the above, every $U_i \cap V_j$ is of dimension smaller than d'. We thus have $\dim(U \cap V) = \max_{i,j} \{\dim(U_i \cap V_j\} < d'$. $\qquad\square$

We now list some useful definitions. A *curve* is a variety with all of its components of dimension one (note that this indeed extends the definition of a curve that was given in Chapter 2 for the special case of $\mathbb{R}^2$). A *k-flat* is a translation of a k-dimensional linear space. A *hypersurface* in $\mathbb{R}^d$ is a variety with all of its components of dimension $d - 1$. Similarly, a hyperplane in $\mathbb{R}^d$ is a $(d-1)$-flat, a hypersphere is a $(d-1)$-dimensional sphere, etc.

As we will see rather often, hypersurfaces are easier to study than lower dimensional varieties. The following lemma is one of the main reasons for this. We will prove this lemma in Chapter ???, after studying more properties of polynomials.

Lemma 4.5. *For every hypersurface $U \subset \mathbb{R}^d$ there exists a polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ such that $\langle f \rangle = \mathbf{I}(U)$.*

Given an irreducible variety $U \subset \mathbb{R}^d$ of dimension d', one might expect U to look like a d'-dimensional set in a small neighborhood around any point $p \in U$. To see that this is not the case, consider the cubic curve $\mathbf{V}(y^2 - x^3 + x^2)$. Even though this is an irreducible variety, it consists of a one-dimensional curve together with the origin; see the left part of Figure 4.1. If we remove the origin, we obtain a set that is not a variety. A similar example is the *Whitney umbrella* $\mathbf{V}(x^2 - y^2 z) \subset \mathbb{R}^3$, which consists of a two-dimensional surface together with the z-axis; see the right part of Figure 4.1. As before, if we remove the line we obtain a set that is not a variety.
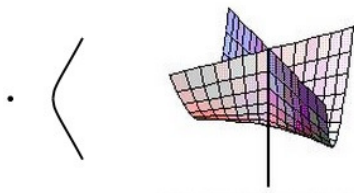
Figure 4.1: A cubic curve in $\mathbb{R}^2$ and the Whitney umbrella.

4.3   Singular points

Consider a variety $U \subset \mathbb{R}^d$ and a point $p \in U$. The *tangent space* of U at p, denoted $T_p U$, is a real vector space in $\mathbb{R}^d$ that contains every vector that has a direction at which one can tangentially pass through p; note that $T_p U$ is not necessarily incident to p.

In Chapter 2 we discussed singular points of curves in $\mathbb{R}^2$. We now extend this definition to singular points of varieties in $\mathbb{R}^d$. Consider a variety $U \subset \mathbb{R}^d$ and a point $p \in U$. Intuitively (and with some exceptions), p is a *singular point* of U if one of the following holds:

- The tangent space $T_p U$ is not well defined. For example, the apex of a circular conical surface is a singular point (see Figure 4.2).
- The point p is contained in more than one irreducible component of U. For example, consider the union of two spheres in $\mathbb{R}^3$ that intersect at a point p and have the same tangent plane at p. Although one might say that the tangent plane is well defined at p, we still consider p as a singular point.
- The dimension of the tangent space $T_p U$ is smaller than the dimension of an irreducible component that contains p. For example, the line of the Whitney umbrella consists of singular points.

While a point that falls under at least one of these three cases is always singular, there are singular points that do not fit any of these cases. For an example, see Section 2.2. A point of U that is not singular is said to be a *regular* point of U.
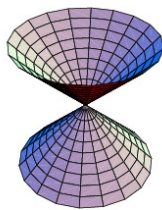
Figure 4.2: The apex of a conical surface is a singular point.

For a more rigorous definition of a singular point, we begin with the special case of hypersurfaces. Given a polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$, the *gradient* of f is

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_d} \right).$$

Notice that this is a generalization of the definition that was given in Chapter 2.

Let $U \subset \mathbb{R}^d$ be a hypersurface. By Lemma 4.5 there exists $f \in \mathbb{R}[x_1, \ldots, x_d]$ such that $\mathbf{I}(U) = \langle f \rangle$. A point $p \in U$ is singular if and only if $\nabla f(p) = 0$ (that is, the gradient is a vector of d zeros). We denote the set of singular points of U as U_{sing}, and the set of regular points of U as U_{reg}. For a discussion about why the condition $\mathbf{I}(U) = \langle f \rangle$ is necessary, see Section 2.2.

Recall that a polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ is said to be *square-free* if in the factorization of f into irreducible factors, no factor has a multiplicity larger than one. Let $f \in \mathbb{R}[x_1, \ldots, x_d]$ be a square-free polynomial and let g be a component of f that depends on x_i. Then $\frac{\partial f}{\partial x_i}$ is not divisible by g, as explained in Section 2.2.

Claim 4.6. *Every hypersurface $U \subset \mathbb{R}^d$ contains a regular point.*

Proof. Let C be a component of U. By Lemma 4.5 there exists $f \in \mathbb{R}[x_1, \ldots, x_d]$ such that $\mathbf{I}(C) = \langle f \rangle$. We assume, without loss of generality, that $f_1 = \frac{\partial f}{\partial x_1}$ is not identically zero. Since f is square-free by definition, the variety $\mathbf{V}(f_1)$ has no common components with C. By Claim 4.4, the intersection $C \cap \mathbf{V}(f_1)$ is of dimension at most $d - 2$. Similarly, the intersection of C with each of the other components of U is of dimension at most $d - 2$. Any point of C that is not contained in $\mathbf{V}(f_1)$ or in any of the other components of U is a regular point. Since a finite number of varieties of dimension at most $d - 2$ cannot cover a variety of dimension $d - 1$, we conclude that C contains regular points. $\qquad\square$

To define singular points of varieties that are not hypersurfaces, we require a generalization of the gradient. Given $f_1, \ldots, f_k \in \mathbb{R}[x_1, \ldots, x_d]$, the *Jacobian matrix* of $f_1, \ldots, f_k$ is

$$
\mathbf{J}_{f_1, \ldots, f_k} = \begin{pmatrix}
\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_d} \\
\frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_d} \\
\cdots & \cdots & \cdots & \cdots \\
\frac{\partial f_k}{\partial x_1} & \frac{\partial f_k}{\partial x_2} & \cdots & \frac{\partial f_k}{\partial x_d}
\end{pmatrix}
$$

Let $U \subset \mathbb{R}^d$ be a variety of dimension d', and consider $f_1, \ldots, f_k \in \mathbb{R}[x_1, \ldots, x_d]$ such that $\mathbf{I}(U) = \langle f_1, \ldots, f_k \rangle$. Then $p \in U$ is a singular point if and only if $\mathrm{rank}\,(\mathbf{J}_{f_1, \ldots, f_k}(p)) < d - d'$. Recall that $\mathrm{rank}\,(\mathbf{J}_{f_1, \ldots, f_k}(p)) + \ker\,(\mathbf{J}_{f_1, \ldots, f_k}(p)) = d$. Moreover, the tangent space $T_p U$ is orthogonal to ∇f_i, for every $1 \leq i \leq k$. Thus, we can intuitively think of $d - \mathrm{rank}\,(\mathbf{J}_{f_1, \ldots, f_k}(p))$ as the dimension of $T_p U$.

It is tempting to think that for a variety $U \subset \mathbb{R}^d$ of dimension d', the ideal $\mathbf{I}(U)$ is generated by $d - d'$ polynomials. To see that this is not necessarily the case, consider

the *twisted cubic* $U = \{(x, x^2, x^3) \in \mathbb{R}^3 \ : \ x \in \mathbb{R}\}$. This is a one-dimensional variety in $\mathbb{R}^3$, but the ideal $\mathbf{I}(U)$ is generated by the three polynomials $x_1 x_3 - x_2^2$, $x_2 - x_1^2$, and $x_3 - x_1 x_2$.

Claim 4.7. *Every variety $U \subset \mathbb{R}^d$ contains a regular point.*

We do not prove Claim 4.7. For more details, see [12, Section 3.3].

4.4 Degrees

We will sometimes consider a *generic* object. When we say that a generic point of $\mathbb{R}^d$ has property X, we mean that the set of points in $\mathbb{R}^d$ that do not have property X are contained in a proper subvariety (and is thus of measure zero). For example, given a plane $h \subset \mathbb{R}^3$, a generic point of $\mathbb{R}^3$ is not contained in h. We can talk about objects that are not points in a similar manner. For example, describing a circle in $\mathbb{R}^2$ requires three parameters — two coordinates for the center and a radius. We can thus think of the circles of $\mathbb{R}^2$ as the points of a three-dimensional space. When saying that a generic circle in $\mathbb{R}^2$ is not incident to the origin, we mean that the set of circles that are incident to the origin are contained in a proper subvariety of this three-dimensional space.

Defining the *degree* of a real variety is somewhat problematic. In a complex space there is a well-defined notion of degree, with many equivalent definitions. For example, a variety $U \subset \mathbb{C}^d$ of dimension d' has degree k if it intersects a generic $(d-d')$-flat of $\mathbb{C}^d$ in exactly k points. To see that this definition does not make sense over the reals, consider a circle in $\mathbb{R}^2$. Intuitively we expect a circle to have degree 2 (and also by the definition given in Chapter 2), but we cannot claim that a generic line intersects a circle in two points.

By corollary 2.2, every real variety is the zero set of a single polynomial, and it might seem tempting to consider the minimum degree of such a polynomial. Unfortunately this definition also has weird consequences. For example, if $U \subset \mathbb{R}^d$ consists of a single point, then it has degree 2.

This has led to a situation where several non-equivalent definitions of degree are being used in $\mathbb{R}^d$, and some works simply avoid defining this notion. In this book, we often consider the degree as being constant, and then obtain the same asymptotic results for any reasonable definition of degree. For our purposes, we define the degree of a variety $U \subset \mathbb{R}^d$ as

$$\min_{\substack{f_1,\ldots,f_k \in \mathbb{R}[x_1,\ldots,x_d] \\ \mathbf{V}(f_1,\ldots,f_k)=U}} \max_{1 \leq i \leq k} \deg f_i. \tag{4.1}$$

That is, the degree of U is the minimum integer D such that U can be defined by a finite set of polynomials of degree at most D.

Given a variety $U \subset \mathbb{R}^d$ of dimension $d - 1$, by Lemma 4.5 the ideal $\mathbf{I}(U)$ is generated by a single polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$. This implies that f is a minimum-degree polynomial in $\mathbf{I}(U)$, and thus that the degree of U is $\deg f$. In particular, we get that (4.1) generalizes the definition of degree from Chapter 2.

Theorem 4.8. *Let $U \subset \mathbb{R}^d$ be a variety of degree k and dimension d'. Then U_{sing} is a variety of dimension smaller than d' and of degree $O_{k,d}(1)$.*

Proof. By definition, there exist polynomials $f_1, \ldots, f_\ell$ of degree at most k such that $\langle f_1, \ldots, f_\ell \rangle = \mathbf{I}(U)$. We have

$$U_{\text{sing}} = \left\{ p \in U : \operatorname{rank} \left(\mathbf{J}_{f_1, \ldots, f_k}(p) \right) < d - d' \right\}.$$

Notice that $\mathbf{J}_{f_1, \ldots, f_k}(p)$ is of rank smaller than $d - d'$ if and only if every $(d - d') \times (d - d')$ minor of $\mathbf{J}_{f_1, \ldots, f_k}(p)$ is zero. Such a minor is an equation of degree at most $k(d - d')$, so U_{sing} is a variety that is defined by a set of polynomials of degree at most $k(d - d') = O_{d,k}(1)$.

To prove that U_{sing} is of dimension smaller than d', we first consider the case where U is an irreducible hypersurface. By Claim 4.7, U contains a regular point, so U_{sing} is a proper subvariety of U (or empty). Since U is irreducible, U_{sing} must have a smaller dimension.

It remains to prove that U_{sing} is of dimension smaller than d' when U is reducible. By the above, the set of points that are singular points of a specific component are of dimension smaller than d', so it remains to consider points that are singular due to an intersection of at least two components. This is straightforward, since the intersection of two distinct irreducible varieties of dimension d' must be of a smaller dimension. $\square$

One disadvantage of the definition of degree in (4.1) is that it does not generalize Bézout's theorem (Theorem 2.5) in dimensions $d \geq 3$. For example, consider the varieties $W = \mathbf{V}((x_1 - 1)^2(x_1 - 2)^2(x_1 - 3)^2 + (x_2 - 1)^2(x_2 - 2)^2(x_2 - 3)^2) \subset \mathbb{R}^3$ and $U = \mathbf{V}(x_3) \subset \mathbb{R}^3$. Note that W is a set of nine lines and of degree six, and that U is a plane and of degree one. However, $U \cap W$ is a set of nine points, so $|U \cap W| > \deg U \cdot \deg W$. As long as we are careful, the failure of Bézout's theorem in higher dimensions will not be a big issue for us.

The degree of a variety also controls the number of irreducible components.

Lemma 4.9. *Let $U \subset \mathbb{R}^d$ be a variety of degree k. Then the number of irreducible components of U is $O_{d,k}(1)$.*

Proof sketch. Consider first the case where U is a hypersurface. By Lemma 4.5, there exists a polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ of degree k such that $\mathbf{V}(f) = U$. We factor f into irreducible factors $f = f_1 \cdot f_2 \cdots f_\ell$. Note that $\ell \le k$. For every $1 \le j \le \ell$ the variety $\mathbf{V}(f_j)$ is irreducible, since otherwise f_j would not be irreducible. Since each irreducible component of U is equal to at least one of the sets U_j, we get that U has at most k irreducible components.

The case where U is not a hypersurface is more involved. We can first handle the part of U that is a hypersurface as above. To handle the m-dimensional components of U (for any $m \le d - 2$), we can project these components onto an $(m+1)$-flat and consider this flat as $\mathbb{R}^{m+1}$. This turns the m-dimensional components into hypersurfaces, which can be handled as above. The projection step should be done carefully to avoid some issues. We only study projections of varieties in Chapter 7, so for now we do not give the full details of this step.

For a different proof over an algebraically closed field, see [17, Lemma A.4]. $\square$

4.5 Polynomial partitioning in $\mathbb{R}^d$

The polynomial partitioning theorem that was presented in Chapter 3 holds in $\mathbb{R}^d$ for every $d \ge 2$. When using this theorem in dimension $d \ge 3$, it would be helpful to have the following slightly revised variant.

Theorem 4.10 (Zahl [107]). *Let $\mathcal{P}$ be a set of m points in $\mathbb{R}^d$. Then for every $1 < r \le m$, there exists an r-partitioning polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ of degree $O(r)$. Moreover, we may assume that $\mathbf{V}(f)$ is a hypersurface.*

Note that the only new claim in Theorem 4.10 is that the partition can always be a hypersurface. To prove this claim, we cannot simply remove lower dimensional components of $\mathbf{V}(f)$, even though such removals cannot cause cells to merge. The issue is that lower dimensional components of $\mathbf{V}(f)$ may contain many points of $\mathcal{P}$, so removing such components may cause some cells to contain too many points. The proof of the hypersurface property requires several additional tools from algebraic geometry, which we will not discuss here.

When using polynomial partitioning in $\mathbb{R}^2$, we relied on Bézout's theorem (Theorem 2.5) to bound the number of cells that are intersected by a curve. In dimension $d \ge 3$, we instead rely on the following more involved result.

Theorem 4.11 (Barone and Basu [5]). *Let U and W be varieties in $\mathbb{R}^d$ such that W is defined by a single polynomial of degree $k_W \geq 2 \deg U$. Then the number of connected components of $U \setminus W$ is $O_d\left(k_W^{\dim U} \deg U^{d-\dim U}\right)$.*

Finally, we will also rely on the following classical result, which is an extension of Theorem 2.7 to $\mathbb{R}^d$.

Theorem 4.12 (Milnor–Thom). *Let $f_1, \ldots, f_m \in \mathbb{R}[x_1, \ldots, x_d]$ be of degree at most k. Then the number of connected components of $\mathbf{V}(f_1, \ldots, f_m)$ is at most*

$$k(2k-1)^{d-1}.$$

4.6   Exercises

Problem 4.1. Prove that every variety can be decomposed into a finite number of irreducible components, without using the degree of the variety (that is, not as in the proof of Lemma 4.9).

Problem 4.2. (a) Consider two polynomials $f, g \in \mathbb{R}[x_1, x_2, x_3]$ of degrees k_f and k_g, respectively. Prove that if f and g have no common factors then $\mathbf{V}(f) \cap \mathbf{V}(g)$ contains at most $k_f k_g$ lines (hint: consider a generic plane).

(b) Consider two polynomials $f, g \in \mathbb{R}[x_1, \ldots, x_d]$ of degrees k_f and k_g, respectively. Prove that if f and g have no common factors then $\mathbf{V}(f) \cap \mathbf{V}(g)$ contains at most $k_f k_g$ flats of dimension $d-2$.

Problem 4.3. (a) Prove Warren's theorem (Theorem 3.2) using the Milnor–Thom theorem (Theorem 4.12). Hint: Change the polynomial by an ε.

(b) Let V be a variety of degree k in $\mathbb{R}^d$. Use Warren's theorem to prove that V has $O_d(k^d)$ connected components.

Problem 4.4. (a) Show that Warren's theorem (Theorem 3.2) is asymptotically tight. That is, for every $d \geq 2$ and k prove that there exists a polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ of degree k such that $\mathbb{R}^d \setminus \mathbf{V}(f)$ has $\Theta(k^d)$ connected components. You might like to first consider the case of $d = 2$.

(b) Show that the Milnor–Thom theorem (Theorem 4.12) is asymptotically tight. That is, for every $d \geq 2$ and k prove that there exists a polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ such that $\mathbf{V}(f)$ has $\Theta(k^d)$ connected components. You might like to use ideas from Problem 4.3.

Chapter 5

The Joints Problem and Degree Reduction

> *"The situation is a little bit like the spread of a disease in a population. If each member of a population is exposed to many other members of the population, then a fairly small outbreak can become an epidemic."* / Larry Guth explaining a degree reduction argument [49].

In some sense the use of polynomial methods to study incidence problems started with a work of Guth and Katz [50], in which they used polynomial methods to solve two problems. In this chapter we study one of these problems: *the joints problem.* Since this was an early application of the polynomial methods, it is relatively simple and a good warmup for working in dimensions $d \geq 3$. In Section 5.2 we study two additional problems that can be solved using similar polynomial arguments. A large part of this chapter is based on ideas presented in [49].

5.1   The joints problem

Let $\mathcal{L}$ be a set of lines in $\mathbb{R}^3$. A *joint* of $\mathcal{L}$ is a point of $\mathbb{R}^3$ that is incident to three lines $\ell_1, \ell_2, \ell_3 \in \mathcal{L}$ such that no plane contains ℓ_1, ℓ_2, and ℓ_3 (and possibly to additional lines of $\mathcal{L}$). In other words, the directions of ℓ_1, ℓ_2, and ℓ_3 should be linearly independent. The joints problem asks for the maximum number of joints in a set of n lines in $\mathbb{R}^3$.

Claim 5.1. *There exists a set of n lines in $\mathbb{R}^3$ that spans $\Theta(n^{3/2})$ joints.*

Proof. Consider $n/3$ lines in the direction of the x-axis, $n/3$ lines in the direction of

the y-axis, and $n/3$ lines in the direction of the z-axis:

$$\mathcal{L} = \left\{ \mathbf{V}(x - a, y - b) \subset \mathbb{R}^3 \ : \ a, b \in \mathbb{N} \ \text{ and } \ 1 \le a, b \le \sqrt{n/3} \right\}$$
$$\bigcup \left\{ \mathbf{V}(x - a, z - c) \subset \mathbb{R}^3 \ : \ a, c \in \mathbb{N} \ \text{ and } \ 1 \le a, c \le \sqrt{n/3} \right\}$$
$$\bigcup \left\{ \mathbf{V}(y - b, z - c) \subset \mathbb{R}^3 \ : \ b, c \in \mathbb{N} \ \text{ and } \ 1 \le b, c \le \sqrt{n/3} \right\}.$$

The joints spanned by this set are

$$\left\{ (a, b, c) \in \mathbb{N}^3 : 1 \le a, b, c \le \sqrt{n/3} \right\}.$$

$\square$

In the above construction, the lines have structure in the sense that every two lines are either parallel or orthogonal. We can obtain the same number of joints without having such structure.

Claim 5.2. *There exists a set of n lines in $\mathbb{R}^3$ that spans $\Theta(n^{3/2})$ joints, with no two lines being parallel or orthogonal.*

Proof. Let Π be a set of m generic planes in $\mathbb{R}^3$, for a parameter m that will be set below. By generic planes, we mean that no two planes are parallel, no three intersect in a line, and no four intersect in a point. Set

$$\mathcal{L} = \{ h \cap h' \ : \ h, h' \in \Pi \text{ and } h \ne h' \}.$$

Since no three planes intersect in a line, $\mathcal{L}$ is a set of $\binom{m}{2}$ distinct lines. We may also assume that no two lines of $\mathcal{L}$ are parallel or orthogonal. We fix the value of m such that $|\mathcal{L}| = n$, and note that $m = \Theta(\sqrt{n})$. For any three distinct planes h, h', h'', the three lines $h \cap h'$, $h \cap h''$, and $h' \cap h''$ form a distinct joint. Thus, the number of joints is $\binom{m}{3} = \Theta(n^{3/2})$. $\square$

The joints problem seems to have started as a Discrete Geometry problem (for examlpe, see [19]), but over the years it also attracted the attention of researchers from Harmonic Analysis. Wolff [106] observed a connection between the joints problem and the Kakeya problem. After a sequence of increasingly better bounds, the problem was completely solved by Guth and Katz.

Theorem 5.3 (Guth and Katz [50]). *The maximum number of joints in a set of n lines in $\mathbb{R}^3$ is $\Theta(n^{3/2})$.*

A generalization of the joints problem to $\mathbb{R}^d$ was derived by Kaplan, Sharir, and Shustin [57] and independently by Quilodrán [73].

We already established the lower bound of Theorem 5.3, so it remains to prove a matching upper bound. The polynomial technique in Chapter 3 was based on studying polynomials that partition a point set into "well-behaved" cells. We now use a different polynomial argument, considering polynomials that contain a point set.

Lemma 5.4. *Given a set $\mathcal{P}$ of m points in $\mathbb{R}^d$ and a positive integer D such that $\binom{d+D}{d} > m$, there exists $f \in \mathbb{R}[x_1, \ldots, x_d] \setminus \{0\}$ of degree at most D such that $\mathcal{P} \subset Z(f)$.*

Proof. In the proof of the polynomial ham sandwich theorem (Theorem 3.6), we argued that the number of distinct monomials in $\mathbb{R}[x_1, \ldots, x_d]$ of degree at most D is $\binom{D+d}{d}$. Set $k = \binom{D+d}{d}$ and note that by assumption $k > m$.

Consider a polynomial f of degree D, and denote the coefficients of the monomials of f as $c_1, \ldots, c_k$. Asking for f to vanish on a point $p \in \mathcal{P}$ corresponds to a linear homogeneous equation in $c_1, \ldots, c_k$. Thus, we have a set of m linear homogenous equations in k variables. Since $k > m$, the system must have a non-trivial solution. A solution corresponds to a choice of coefficients for f such that f vanishes on $\mathcal{P}$, and a non-trivial solution corresponds to a nonzero polynomial $\qquad\square$

Lemma 5.5. *Let $\mathcal{L}$ be a set of lines in $\mathbb{R}^3$ and let $\mathcal{J}$ be the set of joints of $\mathcal{L}$. Then there exists a line of $\mathcal{L}$ that is incident to at most $3|\mathcal{J}|^{1/3}$ of the joints.*

Proof. Assume for contradiction that every line of $\mathcal{L}$ is incident to more than $3|\mathcal{J}|^{1/3}$ joints. Let $f \in \mathbb{R}[x_1, x_2, x_3]$ be a minimum degree nonzero polynomial that vanishes on $\mathcal{J}$. Since $\binom{3+3|\mathcal{J}|^{1/3}}{3} > |\mathcal{J}|$, Lemma 5.4 implies that $\deg f \le 3|\mathcal{J}|^{1/3}$.

Consider a line $\ell \in \mathcal{L}$ and a generic plane h that contains ℓ. Notice that $\gamma = \mathbf{V}(f) \cap h$ is a variety of dimension at most one and of degree at most $\deg f \le 3|\mathcal{J}|^{1/3}$. By applying Bézout's theorem (Theorem 2.5) in the plane h, we get that either γ contains ℓ or $|\gamma \cap \ell| \le 3|\mathcal{J}|^{1/3}$. By assumption ℓ contains more than $3|\mathcal{J}|^{1/3}$ points, so we must have $\ell \subseteq \gamma \subset \mathbf{V}(f)$. That is, $\mathbf{V}(f)$ contains every line of $\mathcal{L}$.

Consider a point $p \in \mathcal{J}$, and let $\ell_1, \ell_2, \ell_3 \in \mathcal{L}$ be three lines that are incident to p and are not contained in a common plane. By the above, we have $\ell_1, \ell_2, \ell_3 \subset \mathbf{V}(f)$. Let $\ell'_1, \ell'_2,$ and ℓ'_3 be the respective translations of $\ell_1, \ell_2,$ and ℓ_3 such that $\ell'_1 \cap \ell'_2 \cap \ell'_3$ is the origin. If p is a regular point of $\mathbf{V}(f)$, then the tangent plane $T_p\mathbf{V}(f)$ must contain $\ell'_1, \ell'_2, \ell'_3$. This is a contradiction since these three lines cannot be contained in a common plane. Thus, p is a singular point of f, which in turn implies that $\nabla f(p) = 0$ for every $p \in \mathcal{J}$.

Without loss of generality, we assume that f involves the coordinate x_1. This implies that the first partial derivative $f_1 = \frac{\partial f}{\partial x_1}$ is not identically zero. By the above property of ∇f, we have that f_1 vanishes on every point of $\mathcal{J}$. This contradicts f being a minimum degree polynomial that vanishes on $\mathcal{J}$, and completes the proof of the lemma. $\square$

After deriving Lemma 5.5, it is straightforward to prove Theorem 5.3.

Proof of Theorem 5.3. Let $\mathcal{L}$ be a set of n lines in $\mathbb{R}^3$, let $\mathcal{J}$ be the set of joints of $\mathcal{L}$, and put $x = |\mathcal{J}|$. We repeatedly consider a line that is incident to at most $3x^{1/3}$ joints of $\mathcal{J}$, remove this line from $\mathcal{L}$, and update $\mathcal{J}$ accordingly (the value of x remains fixed during this process). By Lemma 5.5, such a line exists at every step. Since every line removal destroys at most $3x^{1/3}$ joints, and since after removing all of the lines no joint remains, we have

$$x \le n \cdot 3x^{1/3}.$$

The assertion of the theorem is obtained by tidying up this equation. $\square$

5.2   Additional applications of the polynomial argument

In this section we study two additional uses of the polynomial technique that we used to solved the joints problem.

Reguli.   Let ℓ_1, ℓ_2, ℓ_3 be parallel lines in $\mathbb{R}^3$ and consider the union of all lines in $\mathbb{R}^3$ that intersect ℓ_1, ℓ_2, and ℓ_3. When ℓ_1, ℓ_2, ℓ_3 are contained in a common plane h, every line that intersects them is also contained in h and the union of these lines is h. When ℓ_1, ℓ_2, ℓ_3 are not contained in a common plane no line intersects all three, so the union is empty. You might like to spend a minute thinking about the case where only two of the lines are parallel.

A *regulus* is the union of all lines that intersect three *pairwise-skew* lines ℓ_1, ℓ_2, ℓ_3 in $\mathbb{R}^3$. That is, no two lines of ℓ_1, ℓ_2, ℓ_3 are parallel and no two intersect. One example of a regulus is the hyperbolic paraboloid $\mathbf{V}(z - xy) \subset \mathbb{R}^3$ (see Figure 5.1). Note that this paraboloid contains every line of the form $\mathbf{V}(x - c, z - cy)$ where $c \in \mathbb{R}$, and also every line of the form $\mathbf{V}(y - c, z - cx)$. Lines from the same family of lines are pairwise-skew. The paraboloid can be defined by fixing three lines ℓ_1, ℓ_2, ℓ_3 from the same family and taking the union of all lines that intersect ℓ_1, ℓ_2, ℓ_3.
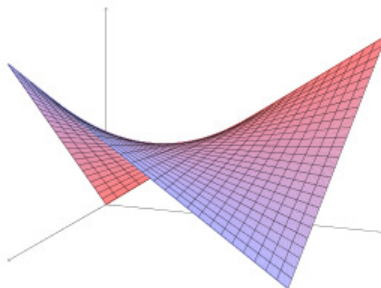
Figure 5.1: The hyperbolic paraboloid $\mathbf{V}(z - xy)$ contains two families of pairwise-skew lines.

Lemma 5.6. *Every regulus in $\mathbb{R}^3$ is contained in an irreducible variety of dimension two and degree two.*

Proof. Consider a regulus S defined by the three pairwise-skew lines ℓ_1, ℓ_2, ℓ_3. Let $\mathcal{P}$ be a set of nine points obtained by arbitrarily choosing three points out of each of the lines ℓ_1, ℓ_2, ℓ_3. Since $\binom{3+2}{3} = 10 > 9$, by Lemma 5.4 there exists a nontrivial polynomial $f \in \mathbb{R}[x, y, z]$ of degree at most two that vanishes on $\mathcal{P}$. By Bézout's theorem, since ℓ_1 intersects $\mathbf{V}(f)$ in at least three points, we have $\ell_1 \subset \mathbf{V}(f)$ (as before, we apply Bézout in a generic plane containing ℓ_1). For the same reason, we also have $\ell_2, \ell_3 \subset \mathbf{V}(f)$. If $\deg f = 1$, then ℓ_1, ℓ_2, ℓ_3 are contained in a common plane, contradicting these lines being pairwise-skew. Similarly, $\mathbf{V}(f)$ cannot be the union of two planes since then at least two of the lines ℓ_1, ℓ_2, ℓ_3 would be on a common plane. We conclude that $\mathbf{V}(f)$ is an irreducible variety of degree two.

Consider a line ℓ' that intersects all three lines ℓ_1, ℓ_2, ℓ_3. The three intersection points are distinct since ℓ_1, ℓ_2, ℓ_3 are disjoint. Since ℓ' intersects $\mathbf{V}(f)$ in at least three points, by Bézout's theorem $\ell' \subset \mathbf{V}(f)$. Since S is the union of all lines that intersect ℓ_1, ℓ_2, ℓ_3, we get that $S \subset \mathbf{V}(f)$. $\square$

With some more work, one can show that the reguli in $\mathbb{R}^3$ are exactly the hyperbolic paraboloids and the hyperboloids of one sheet. However, proving this will not be helpful for practicing polynomial methods, which is the purpose of this section.

Degree reduction. In the proof of Lemma 5.5 we forced lines to be in a variety $\mathbf{V}(f)$ by showing that every line has many common points with $\mathbf{V}(f)$ and applying Bézout's theorem. We then used the same idea in the proof of Lemma 5.6. The following claim is a straightforward generalization of this idea.

Claim 5.7. *Let $\mathcal{L}$ be a set of n lines in $\mathbb{R}^3$. Then there exists a nontrivial polynomial in $\mathbb{R}[x_1, x_2, x_3]$ of degree smaller than $3\sqrt{n}$ that vanishes on all the lines of $\mathcal{L}$.*

Proof. We create a point set $\mathcal{P}$ by arbitrarily choosing $4\sqrt{n}$ points from every line of $\mathcal{L}$. Note that $|\mathcal{P}| \leq 4n^{3/2}$. Since $\binom{3+3\sqrt{n}}{3} > 4n^{3/2}$, Lemma 5.4 implies the existence of a nontrivial polynomial $f \in \mathbb{R}[x_1, x_2, x_3]$ of degree at most $3\sqrt{n}$ that vanishes on $\mathcal{P}$. For every line $\ell \in \mathcal{L}$, since f vanishes on at least $4\sqrt{n}$ points of ℓ, Bézout's theorem implies that ℓ is contained in $\mathbf{V}(f)$. $\square$

When every line of $\mathcal{L}$ contains many intersection points with other lines of $\mathcal{L}$, we can improve the bound of Claim 5.7. This idea is called *degree reduction*, and the quote at the beginning of this chapter refers to its proof.

Lemma 5.8. *Let $\mathcal{L}$ be a set of n lines in $\mathbb{R}^3$, such that each line of $\mathcal{L}$ contains at least k distinct points where it intersects other lines of $\mathcal{L}$ (k may depend on n). Then there exists a nontrivial polynomial in $\mathbb{R}[x_1, x_2, x_3]$ of degree at most $O\left(n^{3/4}/\sqrt{k}\right)$ that vanishes on all the lines of $\mathcal{L}$.*

Proof. When $k \leq 50^2\sqrt{n}$, the lemma is immediately implied by Claim 5.7. We thus assume that $k > 50^2\sqrt{n}$. When n is smaller than some constant, the lemma is immediate by taking a large constant in the $O(\cdot)$-notation. We thus assume that n is at least some sufficiently large constant.

We set a probability $p = 100\sqrt{n}/k$, and consider a subset $\mathcal{L}' \subset \mathcal{L}$ by choosing every line of $\mathcal{L}$ with probability p. With positive probability, $|\mathcal{L}'| < 200n^{3/2}/k$ and every line of $\mathcal{L} \setminus \mathcal{L}'$ has at least $\sqrt{n}$ intersection points with lines of $\mathcal{L}'$. The full details of this standard probabilistic calculation can be found in Section 5.3 below. Since this scenario occurs with positive probability, there exists a subset $\mathcal{L}'$ that satisfies these properties. We consider such a subset.

By Claim 5.7, there exists a nontrivial polynomial $f \in \mathbb{R}[x_1, x_2, x_3]$ of degree at most $3 \cdot \sqrt{200}n^{3/4}/\sqrt{k} < 45n^{3/4}/\sqrt{k}$ that vanishes on every line of $\mathcal{L}'$. Consider a line $\ell \in \mathcal{L} \setminus \mathcal{L}'$. Since $k > 50^2\sqrt{n}$, we have that $\deg f < \sqrt{n}$. Since f vanishes on at least $\sqrt{n}$ points of ℓ, Bézout's theorem implies that ℓ is contained in $\mathbf{V}(f)$. We conclude that $\mathbf{V}(f)$ contains every line of $\mathcal{L}$. $\square$

5.3   (Optional) The probabilistic argument

In this section we prove the probabilistic statement that was made in the proof of Lemma 5.8. We begin by recalling some basic probability. A random variable with a *binomial distribution* $B(n, p)$ represents the number of coin flips that landed on heads when performing n independent coin flips, each with a probability of p for landing heads. For a proof of the following lemma, see for example [2, Theorem A.1.15].

Lemma 5.9 (Chernoff bounds). *Let $X \sim B(n, p)$ where $0 < p < 1$, and let $\delta > 0$. Then*

$$\Pr[X \geq (1 + \delta)np] \leq \left(\frac{e^\delta}{(1 + \delta)^{1+\delta}}\right)^{np}.$$

$$\Pr[X \leq (1 - \delta)np] \leq e^{-pn\delta^2/2}.$$

We are now ready to prove the asserted probabilistic result.

Lemma 5.10. *Let $\mathcal{L}$ be a set of n lines, for a sufficiently large n. Each line contains at least k distinct intersection points with other lines of $\mathcal{L}$, where $50^2\sqrt{n} < k < n$. Let $p = 100\sqrt{n}/k$ and let $\mathcal{L}'$ be a subset of $\mathcal{L}$ obtained by choosing every line of $\mathcal{L}$ with probability p. Then with positive probability $|\mathcal{L}'| < 200n^{3/2}/k$ and every line of $\mathcal{L} \setminus \mathcal{L}'$ contains at least $\sqrt{n}$ distinct intersection points with lines of $\mathcal{L}'$.*

Proof. Note that $|\mathcal{L}'| \sim B(n, p)$. By Lemma 5.9 with $\delta = 1$,

$$\Pr\left[|\mathcal{L}'| \geq 200n^{3/2}/k\right] \leq \left(\frac{e}{4}\right)^{100n^{3/2}/k} < \left(\frac{3}{4}\right)^{100\sqrt{n}}.$$

For a line $\ell \in \mathcal{L}$, we assume that ℓ has exactly k distinct intersection points with the other lines of $\mathcal{L}$, by discarding some intersection points if necessary. Out of these k intersection points, let X_ℓ denote the number of intersection points that ℓ has with lines of $\mathcal{L}'$. Note that $X_\ell \sim B(k, p)$. By Lemma 5.9 with $\delta = 1/2$, we have

$$\Pr\left[X_\ell < \sqrt{n}\right] < \Pr\left[X_\ell \leq 50\sqrt{n}\right] = \Pr\left[X_\ell \leq kp/2\right] \leq e^{-100\sqrt{n}/8} < e^{-10\sqrt{n}}.$$

Recall the union bound principle, stating that the probability of at least one out of a set of events happening is at most the sum of the probabilities of the individual events. In our case, the probability that $|\mathcal{L}'| \geq 200n^{3/2}/k$ or that at least one line of $\mathcal{L} \setminus \mathcal{L}'$ has fewer than $\sqrt{n}$ intersection points with the lines of $\mathcal{L}'$ is smaller than

$$\left(\frac{3}{4}\right)^{100\sqrt{n}} + n \cdot e^{-10\sqrt{n}}.$$

When n is sufficiently large, this probability is smaller than 0.5. Thus, with positive probability $|\mathcal{L}'| < 200n^{3/2}/k$ and every line of $\mathcal{L} \setminus \mathcal{L}'$ has at least $\sqrt{n}$ intersection points with the lines of $\mathcal{L}'$. $\qquad \square$

5.4   Exercises

Problem 5.1. How will the restriction on D in Lemma 5.4 change when:
(a) In every monomial of f, the degree of each variable must be even.
(b) The degree of f must be even.
(c) In every monomial of f, the degree of x_1 must be at least one.

Problem 5.2. Let $\mathcal{L}$ be a set of n lines in $\mathbb{R}^3$. Prove that there exists a nonzero polynomial of degree $O(n^{1/2})$ that vanishes on every line of $\mathcal{L}$. Hint: One way is to apply Lemma 5.4 without thinking about the proof of this lemma.

Problem 5.3.
(a) Let $\mathcal{P}$ be a set of m points in $\mathbb{R}^2$. Use Theorem 5.4 to prove that the number of ellipses in $\mathbb{R}^2$ that are incident to more than $4\sqrt{m}$ points of $\mathcal{P}$ is $O(\sqrt{m})$.
(b) For any $0 < \alpha < 1/2$, use part (a) to prove that the number of ellipses that contain $\Omega(m^{\alpha+1/2})$ points of $\mathcal{P}$ is $O(n^{1/2-\alpha})$ (hint: this requires familiarity with the probabilistic method).

Problem 5.4. Let $\mathcal{L}$ be a set of lines in $\mathbb{R}^d$. A *joint* of $\mathcal{L}$ is a point of $\mathbb{R}^d$ that is incident to d lines $\ell_1, \ldots, \ell_d \in \mathcal{L}$ such that no hyperplane contains $\ell_1, \ldots, \ell_d$. Derive an upper bound for the maximum number of joints of $\mathcal{L}$, and prove that this bound is asymptotically tight.

Problem 5.5. A planar curve in $\mathbb{R}^3$ is a curve contained in some plane in $\mathbb{R}^3$.
(a) Prove Claim 5.7 when the lines are replaced with irreducible planar curves of degree d in $\mathbb{R}^3$. The bound of the claim should change from $3\sqrt{n}$ to $O(\sqrt{n} \cdot d^{3/2})$.
(b) Prove Lemma 5.8 when the lines are replaced with irreducible planar curves of degree d in $\mathbb{R}^3$. The bound of the lemma should change from $O(n^{3/4}/\sqrt{k})$ to $O(n^{3/4}d^{11/4}/\sqrt{k})$. If you prefer not to delve into the probabilistic calculations from Section 5.3, you can instead assume that everything random behaves like its expectation.

Chapter 6

Polynomial Methods in Finite Fields

In this chapter, we use polynomial methods to study incidence-related problems in spaces over finite fields. We focus on two breakthroughs: A solution to the finite field Kakeya problem and the cap set problem. The proofs of these results are short, elegant, and require mostly elementary tools. In Chapter 13 we study point–line incidences in spaces over finite fields, which requires more involved arguments.

6.1   Preliminaries

Recall that a field is *finite* if it contains finitely many elements, and that the *order* of a finite field is the number of elements in it. There is a finite field of order $q \in \mathbb{N} \setminus \{0\}$ if and only if $q = p^r$ for some prime p and positive integer r. Moreover, for every such q there is a unique field of that order (up to an isomorphism). We denote this finite field as $\mathbb{F}_q$ or $\mathbb{F}_{p^r}$.

For a prime p, the field $\mathbb{F}_p$ is the set of integers $\{0, 1, \ldots, p-1\}$ under addition and multiplication mod p. In general, for a prime p and positive integer r, we can define the field $\mathbb{F}_{p^r}$ by considering an irreducible polynomial $f \in \mathbb{F}_p[x]$ of degree r. We can then think of $\mathbb{F}_{p^r}$ as the set of polynomials in $\mathbb{F}_p[x]$ under addition and multiplication mod f. That is, when multiplying two polynomials in $\mathbb{F}_p[x]$ we first perform the standard polynomial multiplication, then replace each coefficient with its value mod p, and finally divide by f and take the remainder. For example, by setting $f = x^2 + 1$ we get $\mathbb{F}_9 = \{0, 1, 2, x, x+1, x+2, 2x, 2x+1, 2x+2\}$. The multiplicative group of $\mathbb{F}_{p^r}$ is cyclic, and the additive group of $\mathbb{F}_{p^r}$ is the direct product of r cyclic groups of

order p.
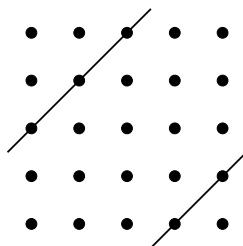
Figure 6.1: The line in $\mathbb{F}_5^2$ defined by $y = x + 2$.

In this chapter we work in the vector space $\mathbb{F}_q^d$, for some prime power q and integer $d \geq 2$. We borrow most of the standard geometric notation from $\mathbb{R}^d$. For example, we refer to $\mathbb{F}_q^2$ as a finite plane and to an element of $\mathbb{F}_q^2$ as a point in this plane. A line in $\mathbb{F}_q^2$ is defined as the zero set of a linear polynomial in $\mathbb{F}_q[x, y]$. Figure 6.1 demonstrates that such a line might have a somewhat surprising behavior.

When working in $\mathbb{F}_{p^r}^d$, our calculations are the aforementioned operation involving mod q and mod an irreducible polynomial of degree r. To distinguish this operation from standard additions and multiplications, we use the notation $x \equiv y$ (for brevity we do not add the mod part afterwards).

We denote by 0_d a vector of d zeros. In other words, it is the origin of a d-dimensional space.

6.2   The finite field Kakeya problem

A *Kakeya set* is a set in $\mathbb{R}^d$ that contain a segment of length 1 in every possible direction. Surprisingly, there exist Kakeya sets in $\mathbb{R}^d$ of measure zero, for any $d \geq 2$. This led to studying the minimum dimension that Kakeya sets must have, which have become a main problem in harmonic analysis. Since this difficult problem is far from being resolved, Wolff [106] suggested that it might be worth considering a finite field variant of the problem. Dvir [26] settled the finite field problem by using a simple and elegant algebraic approach. Some people consider this work as the beginning of the new algebraic techniques in Discrete Geometry.

Consider a finite field $\mathbb{F}_q$, and let ℓ be a line in $\mathbb{F}_q^d$. We can write $\ell = \{u + xv : x \in \mathbb{F}_q\}$, where $u \in \mathbb{F}_q^d$ is a point incident to ℓ and v is the direction vector of ℓ. A point set $\mathcal{P} \subset \mathbb{F}_q^d$ is a *Kakeya set* if it contains at least one line in every direction (that is, at least one line for every direction vector v as defined above). By containing a line, we mean containing the set of points of $\mathbb{F}_q^d$ that are incident to the line. For example,

it is not difficult to verify that in $\mathbb{F}_q^2$ a Kakeya set needs to contain $q+1$ lines with distinct directions.

Theorem 6.1 (Finite field Kakeya [26]). *Let $\mathcal{P} \subset \mathbb{F}_q^d$ be a Kakeya set. Then $|\mathcal{P}| = \Theta_d(q^d)$.*

Theorem 6.1 states that any Kakeya set in $\mathbb{F}_q^d$ contains at least some constant fraction of the points of $\mathbb{F}_q^d$.

It is clear that the only polynomial of $\mathbb{R}[x_1, \ldots, x_d]$ that vanishes on all of $\mathbb{R}^d$ is 0. When working in $\mathbb{F}_q[x_1, \ldots, x_d]$, this claim is no longer true. As a simple example, note that $x_1^2 - x_1 \in \mathbb{F}_2[x_1, \ldots, x_d]$ vanishes on $\mathbb{F}_2^d$. We rely on the following result to address this issue (see Problem 6.1 for a proof).

Lemma 6.2 (Schwartz–Zippel [82, 111]). *Let $f \in \mathbb{F}_q[x_1, \ldots, x_d] \setminus \{0\}$ be of degree k. Then f vanishes on at most kq^{d-1} points of $\mathbb{F}_q^d$.*

We will also rely on the following lemma, which is a finite field variant of Lemma 5.4. The proof is identical to the proof of Lemma 5.4, so we do not repeat it here.

Lemma 6.3. *Given a set $\mathcal{P}$ of m points in $\mathbb{F}_q^d$ and a positive integer k such that $\binom{d+k}{d} > m$, there exists a polynomial $f \in \mathbb{F}_q[x_1, \ldots, x_d] \setminus \{0\}$ of degree at most k such that f vanishes on every point of $\mathcal{P}$.*

By Lemma 6.2, when applying Lemma 6.3 with $k \leq q-1$ the resulting polynomial f does not vanish on all of $\mathbb{F}_q^d$.

Proof of Theorem 6.1. Assume for contradiction that there exists a Kakeya set $\mathcal{P} \subset \mathbb{F}_q^d$ such that $|\mathcal{P}| \leq (q-1)^d/d!$. By Lemma 6.3 there exists a polynomial $f \in \mathbb{F}_q[x_1, \ldots, x_d]$ of degree $1 \leq k \leq q-1$ that vanishes on $\mathcal{P}$. Write $f = \sum_{j=0}^k f_j$, where f_j is a homogeneous polynomial of degree j. By the definition of k we have that $f_k \not\equiv 0$.

Let $v \in \mathbb{F}_q^d \setminus \{0_d\}$ be an arbitrary vector. Since $\mathcal{P}$ is a Kakeya set, it contains a line of direction v. That is, there exists $u \in \mathcal{P}$ such that $\{u + tv : t \in \mathbb{F}_q\} \subset \mathcal{P}$. Since f vanishes on $\mathcal{P}$, we get that $f(u + tv)$ vanishes on every $t \in \mathbb{F}_q$. Since f is of degree at most $q-1$, Lemma 6.2 implies $f(u+tv) \equiv 0$. In particular, the coefficient of t^k in $f(u+tv)$ is zero. Note that the coefficient of t^k in $f(u+tv)$ is $f_k(v)$. We obtain that $f_k(v) \equiv 0$ for every nonzero vector v, so Lemma 6.2 implies that $f_k \equiv 0$. This contradicts f being of degree k, and completes the proof. $\square$

6.3   The cap set problem

A lot of effort has been dedicated to studying the dense sets that do not contain any arithmetic progressions. For example, the following result was cited as one of the two reasons for Klaus Roth's Fields Medal.

Theorem 6.4 (Roth [80]). *There exists a constant $c > 0$ such that the following holds for every positive integer n. Any set $A \subset \{1, 2, \ldots, n\}$ with $|A| \geq cn/\lg \lg n$ contains a 3-term arithmetic progression.*

Many further works improved Theorem 6.4, generalized it, and studied variants of it. We now consider a finite fields variant of the problem, called *the cap set problem.*

We say that three distinct points $p, q, r \in \mathbb{F}_3^n$ form a *3-term arithmetic progression* if there exists $d \in \mathbb{F}_3^n$ such that $q = p + d$ and $r = p + 2d$. This is equivalent to $p + r - 2q = 0_n$ and thus to $p + q + r = 0_n$. We say that a set $A \subset \mathbb{F}_3^n$ is a *cap set* if no three points of A form a 3-term arithmetic progression.
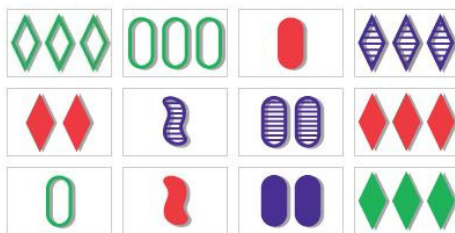
Figure 6.2: The card game *Set* corresponds to $\mathbb{F}_3^4$.

If you know the card game *Set*, you can also think about the cap set problem as such a game. In this game, every card has four properties: number, color, pattern, and shape (see Figure 6.2). The cap set problem is obtained by taking every card to have n properties, and asking for the maximum number of cards that do not contain a set.

It is easy to verify that $\{0, 1\}^n \subset \mathbb{F}_3^n$ is a cap set of size 2^n. In $\mathbb{F}_3^2$ there is no cap set of size larger than four. However, for every $n \geq 3$ there exists a cap set in $\mathbb{F}_3^n$ of size larger than 2^n. Edel [29] proved that there exists a cap set in $\mathbb{F}_3^n$ of size $\Omega(2.217^n)$. The cap set problem asks for the maximum size of a cap set in $\mathbb{F}_3^n$.

By adapting Roth's argument, Meshulam [66] proved that any capset in $\mathbb{F}_3^n$ is of size $O(3^n/n)$. Meshulam's proof is a simple and elegant Fourier transform argument. Bateman and Katz [8] introduced a long and involved analysis that pushed further Meshulam's argument. This sophisticated proof improved the bound to $O(3^n/n^{1+\varepsilon})$, for some small $\varepsilon > 0$.

For years, experts disagreed about whether the answer to the cap set problem should involve a number smaller than 3 at the base of the exponent. Recently, Ellenberg and Gijswijt [34] settled this by proving that every cap set in $\mathbb{F}_3^n$ is of size $O(2.756^n)$. Their proof was surprisingly short and elegant. Rather than using Fourier analysis, it used a somewhat elementary polynomial proof (building on a previous result of Croot, Lev, and Pach [24]).

In Section 6.5 we present an elegant variant of the cap set proof by Tao [98]. Both the original proof and Tao's variant rely on the linear algebra notions of dimension and rank. As a warm-up, in Section 6.4 we solve two other combinatorial problems by using these linear algebra notions.

6.4 Warmups: two distances and odd towns

The two distances problem.   What is the maximum size of a set $\mathcal{P} \subset \mathbb{R}^d$ such that the distance between every two points of the set is 1? By taking the vertices of a d-dimensional simplex with side length 1 in $\mathbb{R}^d$, we obtain $d + 1$ points that span only the distance 1. It is not difficult to show that every set of $d + 2$ points in $\mathbb{R}^d$ spans more than a single distance.

A point set $\mathcal{P}$ is a *two-distance set* if there exist $r, s \in \mathbb{R}$ such that the distance between every pair of points of $\mathcal{P}$ is either r or s. What is the maximum size of a two-distance set in $\mathbb{R}^d$?

Consider the set of all points in $\mathbb{R}^d$ that have two coordinates with value 1 and the other $d - 2$ coordinates with value 0. There are $\binom{d}{2}$ such points, and the distance between every pair of those is either $\sqrt{2}$ or 2. Larman, Rogers, and Seidel [61] showed that the above example is not far from being tight.

Theorem 6.5. *Every two-distance set in $\mathbb{R}^d$ has size at most $\binom{d}{2} + 3d + 2$.*

Proof. Let $\mathcal{P} = \{p_1, p_2, \ldots, p_m\}$ be a two-distance set in $\mathbb{R}^d$, and denote the two distances as $r, s \in \mathbb{R}$. We denote the distances between points $a, b \in \mathbb{R}^d$ as $D(a, b)$. We refer to a point in $\mathbb{R}^d$ as $x = (x_1, \ldots, x_d)$. For $1 \leq j \leq m$, define the polynomial $f_j \in \mathbb{R}[x_1, \ldots, x_d]$ as

$$f_j(x) = \left(D(x, p_j)^2 - r^2 \right) \cdot \left(D(x, p_j)^2 - s^2 \right).$$

Note that f_j vanishes on $q \in \mathbb{R}^d$ if and only if the distance between p_j and q is

either r or s. Every $f_j(x)$ is a linear combination of the polynomials

$$\left(\sum_{j=1}^{d} x_j^2\right)^2, \qquad x_k \sum_{j=1}^{d} x_j^2, \qquad x_k, \qquad x_\ell x_k, \qquad 1,$$

for every $1 \le \ell, k \le d$. Since every f_j is a linear combination of $t = \binom{d}{2} + 3d + 2$ polynomials, we can represent f_j as a vector in $\mathbb{R}^t$. That is, the vector $(v_1, \ldots, v_t)$ corresponds to the polynomial

$$v_1 \cdot \left(\sum_{j=1}^{d} x_j^2\right)^2 + v_2 \cdot x_1 \sum_{j=1}^{d} x_j^2 + v_3 \cdot x_2 \sum_{j=1}^{d} x_j^2 + \cdots + v_t \cdot 1.$$

For $1 \le j \le m$, let V_j be the vector corresponding to $f_j(x)$. Consider $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ such that $\sum_{j=1}^{m} \alpha_j V_j = 0$. Let $g(x) = \sum_{j=1}^{m} \alpha_j f_j(x) = 0$. For some fixed $p_k \in \mathcal{P}$, note that $f_j(p_k) = 0$ for every $j \ne k$ and that $f_j(p_j) = r^2 s^2$. Combining the above implies

$$g(p_k) = \sum_{j=1}^{m} \alpha_j f_j(p_k) = \alpha_k r^2 s^2.$$

Since $g(x) = 0$, we have that $\alpha_k = 0$. That is, the only solution $\sum_{j=1}^{m} \alpha_j V_j = 0$ is $\alpha_1 = \cdots = \alpha_m = 0$. This implies that the vectors $V_1, \ldots, V_m$ are linearly independent. Since these vectors are in $\mathbb{R}^t$, we conclude that $m \le t = \binom{d}{2} + 3d + 2$. $\qquad\square$

In the proof of Theorem 6.5 we think of a space of polynomials as a vector space, and study the dimension of this space. The following problem can be solved in a similar manner. We present the proof in slightly different manner, to involve the rank of a matrix. This will prepare us for the cap sets proof in the following section.

The odd town problem. The odd town problem studies sets of odd size that have even-sized intersections.

Theorem 6.6. *In the town of Liouville there are n people and m clubs. Every club has an odd number of members. For every two clubs, the number of people who are in both is even. Then $m \le n$.*

Proof. Let $P = \{p_1, \ldots, p_n\}$ be the set of people and let $C = \{c_1, \ldots, c_m\}$ be the set of clubs. Let Z be an $m \times m$ matrix with entries in $\mathbb{F}_2$. The value of $Z_{j,k}$ is the number of people who are in both c_j and c_k, modulo 2. In other words, Z has 1's on the main diagonal and 0's everywhere else. This implies that $\mathrm{rk}(Z) = m$.

For $1 \leq j \leq n$, let W^j be an $m \times m$ matrix with entries in $\mathbb{F}_2$, such that $W^j_{k,\ell} = 1$ if and only if person p_j is a member of both c_j and c_k. Clearly $Z = \sum_{j=1}^n W^j$, under addition mod 2.

We now study W^j for a fixed $1 \leq j \leq n$. For $1 \leq k \leq m$, if person p_j is not a member of club c_k then the k'th row of W^j is all zeros. Also, all the nonzero rows of W^j are identical (there are 1's exactly for the clubs of p_j). This implies that $\mathrm{rk}(W^j) \leq 1$. Since Z is a sum of n matrices of rank at most one, we have $\mathrm{rk}(Z) \leq n$. Recalling that $\mathrm{rk}(Z) = m$, we conclude that $m \leq n$. $\qquad\square$

In the two distances problem we studied the dimension of a vector space. In the two distances problem we studied ranks of matrices. For the cap set problem, we will further generalize this approach to three-dimensional objects.

6.5 The slice rank method

We now return to studying sets in $\mathbb{F}_3^n$. Consider a set $A \subset \mathbb{F}_3^n$ and a polynomial $f : A \times A \times A \to \mathbb{F}_3$. We say that $f(x, y, z)$ has a *slice rank* of 1 if we can write

$$f(x, y, z) = g(x) \cdot h(y, z) \quad \text{or} \quad f(x, y, z) = g(y) \cdot h(x, z) \quad \text{or} \quad f(x, y, z) = g(z) \cdot h(x, y),$$

for some polynomials $g : A \to \mathbb{F}_3$ and $h : A \times A \to \mathbb{F}_3$. More generally, we say that $f(x, y, z)$ has a slice rank of k if k is the smallest integer that satisfies the following: There exist k polynomials $f_1, \ldots, f_k : A \times A \times A \to \mathbb{F}_3$ of slice rank 1 such that $f(x, y, z) = \sum_{j=1}^k f_j(x, y, z)$. We write $\mathrm{sr}(f) = k$. As a first observation, note that the slice rank of a polynomial is at most the number of monomials it has. As another example, when $A = \mathbb{F}_3$ we have $\mathrm{sr}(x + y + z) = 2$.

We consider $g, h : A \to \mathbb{F}_3$ as the same function if $g(a) = h(a)$ for every $a \in A$. That is, g and h may be distinct as general functions but identical when their range is restricted to A. The heart of the cap set proof lies in the following property of slice rank.

Lemma 6.7. *Let $A \subset \mathbb{F}_3^n$ and let $f : A \times A \times A \to \mathbb{F}_3$ satisfy that $f(x, y, z) \neq 0$ if and only if $x = y = z$. Then $\mathrm{sr}(f) = |A|$.*

Note the similarity to the odd town proof, where $c_j \cap c_k \not\equiv 0 \pmod 2$ if and only if $j = k$.

Proof. For $a \in A$, we define the function $\mathbf{1}_a : A \to \mathbb{F}_3$ as

$$\mathbf{1}_a(x) = \prod_{j=1}^n \left(1 - (a_j - x_j)^2 \right).$$

For $c, d \in \mathbb{F}_3$, note that $1 - (c - d)^2 \equiv 0$ unless $c = d$. Thus, $\mathbf{1}_a(x) \equiv 1$ when $a = x$ and otherwise $\mathbf{1}_a(x) \equiv 0$.

We have that

$$f(x, y, z) = \sum_{a \in A} \mathbf{1}_a(x) f(a, y, z).$$

Since this is a sum of $|A|$ polynomials of slice rank one, we get that $\mathrm{sr}(f) \le |A|$. Note that this bound holds for every polynomial $f(x, y, z)$.

By the definition of $\mathrm{sr}(f)$, there exist polynomials $f_j : A \to \mathbb{F}_3$ and $g_j : A \times A \to \mathbb{F}_3$ such that

$$f(x, y, z) = \sum_{j=1}^{s} f_j(x) g_j(y, z) + \sum_{j=s+1}^{t} f_j(y) g_j(x, z) + \sum_{j=t+1}^{\mathrm{sr}(f)} f_j(z) g_j(x, y). \tag{6.1}$$

Let P be the space of functions from A to $\mathbb{F}_3$. For any function $h : A \to \mathbb{F}_3$ we have $h(x) = \sum_{a \in A} h(a) \cdot \mathbf{1}_a(x)$. Thus, P is spanned by $|A|$ polynomials. These polynomials are linearly independent, so $\dim P = |A|$.

Let P' be the set of polynomials $h \in P$ that satisfy the following: For every $1 \le j \le s$ we have

$$\sum_{a \in A} f_j(a) h(a) \equiv 0.$$

This is a set of s homogeneous equations in elements of P, so $\dim P' \ge |A| - s$.

For a polynomial $h \in P'$, denote the *support* of h as $S_h = \{a \in A : h(a) \ne 0\}$. Fix $h \in P'$ with a maximal support. If $|S_h| < \dim P'$ then there is a nonzero $h' \in P'$ that vanishes on every point of S_h. (As above, asking a polynomial to vanish on a point corresponds to one homogeneous equation.) Then $h + h' \in P'$ and has a larger support than h, contradicting our choice of h. Then $h + h'$ would have a larger support than h, contradicting our choice of h. We conclude that $|S_h| \ge \dim P' \ge |A| - s$.

Recalling (6.1), we have that

$$F(y, z) = \sum_{a \in A} h(a) \cdot f(a, y, z)$$

$$= \sum_{a \in A} h(a) \left(\sum_{j=1}^{s} f_j(a) g_j(y, z) + \sum_{j=s+1}^{t} f_j(y) g_j(a, z) + \sum_{j=t+1}^{\mathrm{sr}(f)} f_j(z) g_j(a, y) \right)$$

$$= \sum_{j=1}^{s} \left(g_j(y, z) \sum_{a \in A} f_j(a) h(a) \right) + \sum_{j=s+1}^{t} \left(f_j(y) \sum_{a \in A} g_j(a, z) h(a) \right)$$

$$+ \sum_{j=t+1}^{\mathrm{sr}(f)} \left(f_j(z) \sum_{a \in A} g_j(a, y) h(a) \right)$$

By the definition of P', we have that $\sum_{a \in A} f_j(a) h(a) = 0$ and may remove the first term in the above expression. Rewriting the above, there exist polynomials $\varphi_j : A \to \mathbb{F}_3$ such that

$$F(y, z) = \sum_{j=s+1}^{t} f_j(y) \varphi_j(z) + \sum_{j=t+1}^{\mathrm{sr}(f)} f_j(z) \varphi_j(y). \tag{6.2}$$

Imitating the proof of Theorem 6.6, we complete the current proof by double counting the rank of a matrix. Write $A = \{a_1, a_2, \ldots, a_{|A|}\}$. Let Z be an $|A| \times |A|$ matrix such that $Z_{j,k} = F(a_j, a_k)$. By the assumption of the theorem, we have $f(x, y, z) = 0$ whenever $y \neq z$. This implies that $F(y, z) = \sum_{a \in A} h(a) \cdot f(a, y, z) = 0$ whenever $y \neq z$. Thus, Z is a diagonal matrix. Similarly, we get that $F(y, y) = h(y) \cdot f(y, y, y)$. This implies that $F(y, y) \neq 0$ for every $y \in S_h$. Recalling that $|S_h| \geq |A| - s$, we get that $\mathrm{rk}(Z) \geq |A| - s$.

For $s < j \leq t$, let W^j be the $|A| \times |A|$ matrix defined by $W^j_{k,\ell} = f_j(a_k) \cdot \varphi_j(a_\ell)$. For $t < j \leq \mathrm{sr}(f)$, let W^j be the $|A| \times |A|$ matrix defined by $W^j_{k,\ell} = f_j(a_\ell) \cdot \varphi_j(a_k)$. It can be easily verified that every two rows of every W^j are linearly dependent. Thus, the rank of every W^j is at most 1.

By (6.2), we note that $Z = \sum_{j=s+1}^{\mathrm{sr}(f)} W_j$. Since Z is a sum of $\mathrm{sr}(f) - s$ matrices of rank at most one, we get that $\mathrm{rk} Z \leq \mathrm{sr}(f) - s$. Combining this with $\mathrm{rk}(Z) \geq |A| - s$ implies $|A| \leq \mathrm{sr}(f)$, which completes the proof. $\square$

The proof of the cap sets result is now mostly an application of Lemma 6.7.

Theorem 6.8. *Let A be a cap set in $\mathbb{F}_3^n$. Then $|A| = O(2.756^n)$.*

Proof. Since A is a cap set, we have that $a, b, c \in A$ satisfy $a + b + c = 0_n$ if and only if $a = b = c$. Consider the polynomial $f : A \times A \times A \to \mathbb{F}_3$ defined as

$$f(x, y, z) = \prod_{j=1}^{n} \left(1 - (x_j + y_j + z_j)^2\right). \tag{6.3}$$

Similarly to an argument in the proof of Lemma 6.7, we note that $f(x, y, z) \equiv 1$ if $x + y + z = 0_n$ and otherwise $f(x, y, z) \equiv 0$. That is, for $a, b, c \in A$ we have that $f(a, b, c) \equiv 1$ if and only if $a = b = c$. We may thus apply Lemma 6.7 with A and $f(x, y, z)$, to obtain that $\mathrm{sr}(f) = |A|$.

For $x = (x_1, \ldots, x_n)$ and $p = (p_1, \ldots, p_n)$, we write $x^p = x_1^{p_1} x_2^{p_2} \cdots x_n^{p_n}$. We also write $|p| = p_1 + p_2 + \cdots + p_n$ with *summation over* $\mathbb{R}$. We say that degree of a monomial of $f(x, y, z)$ in x is the sum of the degrees of $x_1, \ldots, x_n$ in this monomial (over $\mathbb{R}$). Since $\deg f(x, y, z) \le 2n$, each monomial of $f(x, y, z)$ has degree at most $2n/3$ in at least one of x, y, and z. We can thus rewrite (6.3) as

$$f(x, y, z) = \sum_{\substack{p \in \mathbb{F}_3^n \\ |p| \le 2n/3}} x^p g_{x,p}(y, z) + \sum_{\substack{p \in \mathbb{F}_3^n \\ |p| \le 2n/3}} y^p g_{y,p}(x, z) + \sum_{\substack{p \in \mathbb{F}_3^n \\ |p| \le 2n/3}} z^p g_{z,p}(x, y). \tag{6.4}$$

Here $g_{x,p}, g_{y,p}, g_{z,p} : A \times A \to \mathbb{F}_3$ are polynomials of degree at most $2n$.

Set $r = |\{p \in \mathbb{F}_3^n : |p| \le 2n/3\}|$. By (6.4) we have that $\mathrm{sr}(f) \le 3r$. Recalling that $\mathrm{sr}(f) = |A|$ gives $|A| \le 3r$. It remains to derive an upper bound for r. This can be done using elementary combinatorial tools, and is in some sense less interesting.

Given $p \in \mathbb{F}_3^n$, for $0 \le j \le 2$ we denote by m_j the number of coordinates of p that are equal to j. Using this notation, we get that

$$r = \sum_{\substack{m_0 + m_1 + m_2 = n \\ m_1 + 2m_2 \le 2n/3}} \frac{n!}{m_0! m_1! m_2!}.$$

The multinomial theorem (for example, see [14, Chapter 4]) implies

$$\left(1 + x + x^2\right)^n = \sum_{m_0 + m_1 + m_2 = n} \frac{n!}{m_0! m_1! m_2!} \cdot x^{m_1 + 2m_2}.$$

Assuming that $0 < x < 1$, the above leads to

$$x^{-2n/3} \left(1 + x + x^2\right)^n = \sum_{m_0+m_1+m_2=n} \frac{n!}{m_0!m_1!m_2!} \cdot x^{m_1+2m_2-2n/3}$$

$$> \sum_{\substack{m_0+m_1+m_2=n \\ m_1+2m_2 \leq 2n/3}} \frac{n!}{m_0!m_1!m_2!} \cdot x^{m_1+2m_2-2n/3}$$

$$> \sum_{\substack{m_0+m_1+m_2=n \\ m_1+2m_2 \leq 2n/3}} \frac{n!}{m_0!m_1!m_2!} = r.$$

Set $g(x) = x^{-2/3} \left(1 + x + x^2\right)$. It can be easily verified that the minimum of $g(x)$ when $0 < x < 1$ is obtain for $x = (\sqrt{33} - 1)/8$. This minimum value satisfies $g(x) < 2.76$, implying that $|A| \leq 3r = O(2.76^n)$. $\qquad\square$

6.6 Exercises

Problem 6.1. It is common to prove Lemma 6.2 (Schwartz–Zippel) by induction on d. In this question you will prove Lemma 6.2 in a different way.

Let f_k be the homogeneous component of degree k of f. Let $p \in \mathbb{F}_q^d \setminus \{0_n\}$ such that $f_k(p) \neq 0$. Partition $\mathbb{F}_d^q$ into q^{d-1} lines with direction p. Complete the proof by showing that f is never identically zero when restricted to such a line.

Problem 6.2. For $0 < \alpha < 1$, we say that $\mathcal{P} \subset \mathbb{F}_q^d$ is a q^α-Kakeya set if for every $u \in \mathbb{F}_q^d \setminus \{0_d\}$ there exists a line with direction u that contains at least q^α points of $\mathcal{P}$. Adapt the proof of finite field Kakeya theorem to obtain a lower bound for the minimum size of a q^α-Kakeya set.

Problem 6.3. Construct a two-distance set in $\mathbb{R}^n$ of size $\binom{n}{2} + 1$.

Problem 6.4. A library contains $n + 1$ books and has n members. Every member read at least one book from the library. Prove that there exist two disjoint sets of members that read exactly the same set of books. It does not matter how many people from the same set read a book, and some members may not be in either set.

Problem 6.5. Solve the library problem (Problem 6.4) with $n+2$ books by applying *Radon's Theorem*. Radon's theorem states that any set of $d + 2$ points in $\mathbb{R}^d$ can be partitioned into two disjoint subsets whose convex hulls intersect.

Problem 6.6. In the even town problem every club has an even number of members and every two clubs have an odd number of common members.

(a) Use a simple reduction to the odd town problem, to show that there are at most $n + 1$ clubs.

(b) Assuming that n is even, imitate the proof of Theorem 6.6 to show that there are at most n clubs.

Chapter 7

Constant-degree Polynomial Partitioning and Incidences in $\mathbb{C}^2$

In this chapter we study incidences with complex algebraic curves in the complex plane $\mathbb{C}^2$. Most of the time we will think of $\mathbb{C}^2$ as $\mathbb{R}^4$, so this chapter is also our first step in studying incidences in $\mathbb{R}^d$ (where $d \geq 3$). To handle such incidence problems we introduce an alternative way of using polynomial partitioning, which we refer to as *constant-degree polynomial partitioning*.[1] In $\mathbb{R}^2$ this technique leads to incidence bounds that are slightly worse than those presented in Chapter 3. However, in dimension $d \geq 3$ it is often significantly simpler to derive incidence bounds by using constant-degree polynomial partitioning.

7.1   Introduction: Incidence issues in $\mathbb{C}^2$ and $\mathbb{R}^d$

Consider a point-curve incidence problem in $\mathbb{C}^2$. That is, we have a set of complex points and curves in $\mathbb{C}^2$ where each curve is defined by a polynomial in $\mathbb{C}[z_1, z_2]$ (and the variables z_1 and z_2 get values in $\mathbb{C}$). It is tempting to try to imitate the proof from Chapter 3 for incidences in $\mathbb{R}^2$ — combinatorially obtain a weak incidence bound, partition $\mathbb{C}^2$ using polynomial partitioning, and then apply the weak bound separately in every cell. Unfortunately, partitioning polynomials do not exist in complex spaces, since removing a variety from $\mathbb{C}^d$ cannot disconnect it into more than one connected component. That is, we cannot split the space into cells using varieties.

One way to overcome the above issue is to think of $\mathbb{C}^2$ as $\mathbb{R}^4$, since partitioning

---

[1]As usual, there is no standard name for this technique. Some papers refer to it as "low degree polynomial partitioning".

polynomials do exist in $\mathbb{R}^4$. Specifically, we define four real variables x_1, x_2, y_1, y_2 and rewrite the two complex coordinates as $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$. Let the map $\phi : \mathbb{C}^2 \to \mathbb{R}^4$ be defined as

$$\phi(x_1 + iy_1, x_2 + iy_2) = (x_1, y_1, x_2, y_2).$$

Using ϕ, we move from a point-curve incidence problem in $\mathbb{C}^2$ to an incidence problem with two-dimensional varieties in $\mathbb{R}^4$. Indeed, consider a curve in $\mathbb{C}^2$ that is defined by a polynomial $f \in \mathbb{C}[z_1, z_2]$. Asking f to vanish on a point $(x_1 + iy_1, x_2 + iy_2)$ is equivalent to asking both the real and the imaginary parts of f to vanish on x_1, y_1, x_2, y_2. That is, $\phi(\mathbf{V}(f)) \subset \mathbb{R}^4$ is a variety defined by two polynomials in $\mathbb{R}[x_1, y_1, x_2, y_2]$ of degree at most $\deg f$. Note that ϕ maintains the point-curve incidences.

We now point out a basic issue that arises when studying incidence problems in dimension $d \geq 3$. For simplicity, we consider one of the simplest cases: incidences between points and planes in $\mathbb{R}^3$. Let $\ell \subset \mathbb{R}^3$ be a line, let $\mathcal{P}$ be a set of m points on ℓ, and let Π be a set of n planes that contain ℓ (e.g., see Figure 7.1). This construction satisfies $I(\mathcal{P}, \Pi) = mn$, implying that the problem is trivial.
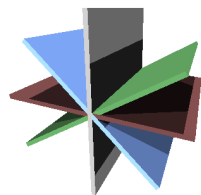
Figure 7.1: By having planes that contain a common line, we can obtain mn point-plane incidences.

To turn the point-plane problem into a non-trivial one, one adds additional restrictions on the points and planes. The most common restriction is probably requiring the incidence graph of $\mathcal{P} \times \Pi$ not to contain a copy of $K_{s,t}$ (for some constants s and t). This problem is interesting (and open), and to obtain some bound for it we try to adapt our proof for the case of point-curve incidences in $\mathbb{R}^2$. Such a restriction arises naturally in many cases. For example, assuming that no k points of $\mathcal{P}$ are collinear implies that the incidence graph contains no $K_{k,2}$. Below we will see that this restriction is also useful when studying incidences in $\mathbb{C}^2$.

Now that we have a non-trivial point-plane incidence problem in $\mathbb{R}^3$, we would like to study it using polynomial partitioning as in Chapter 3. By inspecting the proof of

our weak incidence bound (Lemma 3.4), we notice that it is not only valid for curves in $\mathbb{R}^2$, but rather for any set of varieties in $\mathbb{R}^d$ and also after replacing $\mathbb{R}$ with another field. In particular, this lemma is just a bound on the number of edges in a bipartite graph with no copy of $K_{s,t}$, and does not involve any geometry.

Lemma 7.1. *Let $\mathbb{F}$ be a field and let d be a positive integer. Let $\mathcal{P}$ be a set of m points in $\mathbb{F}^d$ and let $\mathcal{V}$ be a set of n subsets of $\mathbb{F}^d$. If the incidence graph of $\mathcal{P} \times \mathcal{V}$ contains no copy of $K_{s,t}$, then*

$$I(\mathcal{P}, \mathcal{V}) = O_{s,t}\left(mn^{1-\frac{1}{s}} + n\right).$$

The polynomial partitioning theorem also holds in any dimension, and we can use it to partition $\mathcal{P}$. In fact, repeating the analysis for incidences in the cells of the partition as in Chapter 3 is straightforward. On the other hand, the last step of the analysis — handling incidences on the partition itself — becomes significantly more difficult than in the planar case. The partition is a two-dimensional variety, possibly of a large degree. It can contain most of the points of $\mathcal{P}$ and intersect most of the planes of Π in high-degree curves.

When deriving incidence bounds in dimension $d \geq 3$, usually the main difficulty is bounding the number of incidences on the partition. Solymosi and Tao [90] introduced the constant-degree polynomial partitioning technique to overcome this difficulty.

7.2 Curves in higher dimensions

In this brief section we present a generalization of the point-curve incidence bound in $\mathbb{R}^2$ (Theorem 3.3) to higher dimensions. We will rely on this result below, and it is also a good introduction to projections of varieties onto lower dimensional spaces.
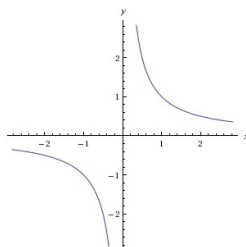
Figure 7.2: The curve $\mathbf{V}(xy - 1)$.

Let $U \subset \mathbb{R}^d$ be a variety of degree k and of dimension d'. Let $\pi : \mathbb{R}^d \to \mathbb{R}^e$ be a standard projection: a linear map that keeps e out of the d coordinates of a point in

$\mathbb{R}^d$. The projection $\pi(U)$ is not necessarily a variety. For example, the projection of $V(xy - 1) \subset \mathbb{R}^2$ onto the x-axis is the set $\{x \in \mathbb{R}^2 : x \neq 0\}$, which is not a variety (see Figure 7.2). However, $\pi(U)$ is *contained* in a variety of dimension at most d' and of degree $O_{k,d}(1)$ (e.g., see [12, Proposition 2.8.6][2]). We denote this variety as $\overline{\pi(U)}$.

Lemma 7.2. *Consider an integer $d \geq 2$. Let $\mathcal{P}$ be a set of m points and let Γ be a set of n varieties of dimension at most one and degree at most k, both in $\mathbb{R}^d$. If the incidence graph of $\mathcal{P} \times \Gamma$ contains no copy of $K_{s,t}$, then*

$$I(\mathcal{P}, \Gamma) = O_{s,t,k,d}\left(m^{\frac{s}{2s-1}} n^{\frac{2s-2}{2s-1}} + m + n\right).$$

Proof. We will assume that no two varieties of Γ have a common component. It is not difficult to handle the case where there are common components (see Problem 3.5). We split every reducible one-dimensional variety of Γ into irreducible curves and possibly also one zero-dimensional variety of degree at most k. By Lemma 4.9, every variety of Γ consists of $O_{k,d}(1)$ irreducible components. Thus, after the splitting step we have $|\Gamma| = O_{k,d}(n)$.

We first bound the number of incidences between $\mathcal{P}$ and the zero-dimensional varieties of Γ. By Theorem 4.12, the number of points in such a variety is $O_{d,k}(1)$. Since each zero-dimensional variety participates in $O_{d,k}(1)$ incidences, in total such varieties contribute $O_{d,k}(n)$ incidences.

We remove the zero-dimensional varieties from Γ, so that Γ becomes a set of $O_{k,d}(n)$ irreducible curves. We perform a generic rotation of $\mathbb{R}^d$ around the origin and then project $\mathcal{P}$ and Γ onto the x_1x_2-plane (that is, we use the projection $\pi(x_1, x_2, \ldots, x_d) = (x_1, x_2)$). This process is equivalent to a projection of $\mathcal{P}$ and Γ onto a generic 2-flat of $\mathbb{R}^d$. We set

$$\mathcal{P}' = \{\pi(p) : p \in \mathcal{P}\}, \quad \text{and} \quad \Gamma' = \{\overline{\pi(\gamma)} : \gamma \in \Gamma\}.$$

Projecting the elements of Γ might introduce new intersection points between the curves of Γ'. However, since we first perform a generic rotation, we may assume that the projection does not lead to new incidences. This implies that $I(\mathcal{P}', \Gamma') = I(\mathcal{P}, \Gamma)$ and that the incidence graph of $\mathcal{P}' \times \Gamma'$ does not contain a copy of $K_{s,t}$. Also due to the generic rotation, we may assume that the points of $\mathcal{P}'$ are distinct and that no two curves of Γ' have a common irreducible component. We apply our point-curve incidence result in $\mathbb{R}^2$ (Theorem 3.3) on $\mathcal{P}'$ and Γ'. This yields the assertion of the lemma. $\qquad \square$

---

[2]I still need to add a proper reference for the degree property.

While Lemma 7.2 is a generalization of Theorem 3.3 to $\mathbb{R}^d$, we can obtain a better bound when $d \geq 3$. For more details see Chapter 11.

7.3 Constant-degree polynomial partitioning

In this section we introduce the constant-degree polynomial partitioning technique. Instead of immediately considering incidences in higher dimensions, we first use this technique to prove a weaker version of the Szemerédi–Trotter theorem. This allows us to see the basic ideas without handling the additional issues that arise in higher dimensions.

Theorem 7.3. *Let $\mathcal{P}$ be a set of m points and let $\mathcal{L}$ be a set of n lines, both in $\mathbb{R}^2$. Then for any $\varepsilon > 0$, we have $I(\mathcal{P}, \mathcal{L}) = O_\varepsilon(m^{2/3+\varepsilon}n^{2/3} + m + n)$.*

A warning for readers who are not used to working with asymptotic notation: It is risky to use an $O(\cdot)$-notation in the statement of the bound that is proved by induction. To see the issue, we now prove the false claim $2^n = O(n)$ by induction on n. For the induction basis, the claim holds for $n \leq 10$ by taking the hidden constant in the $O(\cdot)$-notation to be sufficiently large. For the induction step, assume that $2^n = O(n)$ for a specific value of n and note that $2^{n+1} = 2 \cdot 2^n = 2 \cdot O(n) = O(n)$. The mistake in the above argument is that the induction step does not really close. While we get an asymptotically correct upper bound for 2^{n+1}, the hidden constant that we get in the $O(\cdot)$-notation is larger than the constant from the induction hypothesis. That is why in the following proof we require the α variables.

Proof of Theorem 7.3. We prove the theorem by induction on $m + n$. Specifically, we prove by induction that, for any fixed $\varepsilon > 0$, there exist constants α_1, α_2 such that

$$I(\mathcal{P}, \mathcal{L}) \leq \alpha_1 m^{2/3+\varepsilon} n^{2/3} + \alpha_2(m + n).$$

For the induction basis, the bound holds for small $m + n$ (e.g., for $m + n \leq 100$) by taking α_1 and α_2 to be sufficiently large.

For the induction step, we first recall our weak incidence bound (Lemma 7.1), which implies $I(\mathcal{P}, \mathcal{L}) = O(m\sqrt{n} + n)$. When $m = O(\sqrt{n})$ this implies $I(\mathcal{P}, \mathcal{L}) = O(n)$, which completes the proof. Thus, we may assume that

$$n = O\left(m^2\right). \tag{7.1}$$

We take r to be a sufficiently large constant, whose value depends on ε and will be determined below. Let f be an r-partitioning polynomial of $\mathcal{P}$. According to

the polynomial partitioning theorem (Theorem 3.1), f is of degree $O(r)$ and $\mathbf{V}(f)$ partitions $\mathbb{R}^2$ into connected cells, each containing at most m/r^2 points of $\mathcal{P}$. By Warren's theorem (Theorem 3.2), the number of cells is $c = O(r^2)$. The relations between the constants of this proof are[3]

$$2^\varepsilon \ll r \ll \alpha_2 \ll \alpha_1.$$

Let $\mathcal{L}_0$ denote the subset of lines of $\mathcal{L}$ that are contained in $\mathbf{V}(f)$, and let $\mathcal{P}_0 = \mathcal{P} \cap \mathbf{V}(f)$. Denote the cells of the partition as $C_1, \ldots, C_c$. For $j = 1, \ldots, c$, put $\mathcal{P}_j = \mathcal{P} \cap C_j$ and let $\mathcal{L}_j$ denote the set of lines of $\mathcal{L}$ that intersect C_j. Note that

$$I(\mathcal{P}, \mathcal{L}) = I(\mathcal{P}_0, \mathcal{L}_0) + I(\mathcal{P}_0, \mathcal{L} \setminus \mathcal{L}_0) + \sum_{j=1}^{c} I(\mathcal{P}_j, \mathcal{L}_j).$$

For any line $\ell \in \mathcal{L} \setminus \mathcal{L}_0$, by Bézout's theorem (Theorem 2.5) ℓ and $\mathbf{V}(f)$ have $O(r)$ common points. This immediately implies

$$I(\mathcal{P}_0, \mathcal{L} \setminus \mathcal{L}_0) = O(nr). \tag{7.2}$$

Set $m_0 = |\mathcal{P}_0|$ and $m' = m - m_0$; that is, m' is the number of points of $\mathcal{P}$ that are in the cells. Since f is of degree $O(r)$, we get that $\mathbf{V}(f)$ can contain at most $O(r)$ lines. This implies

$$I(\mathcal{P}_0, \mathcal{L}_0) = O(m_0 r). \tag{7.3}$$

It remains to bound $\sum_{j=1}^{c} I(\mathcal{P}_j, \mathcal{L}_j)$. For $j = 1, \ldots, c$, put $m_j = |\mathcal{P}_j|$ and $n_j = |\mathcal{L}_j|$. Note that $m' = \sum_{j=1}^{c} m_j$, and recall that $m_j \le m/r^2$ for every $1 \le j \le c$. By the induction hypothesis, we have

$$\sum_{j=1}^{c} I(\mathcal{P}_j, \mathcal{L}_j) \le \sum_{j=1}^{c} \left(\alpha_1 m_j^{2/3+\varepsilon} n_j^{2/3} + \alpha_2 (m_j + n_j) \right)$$

$$\le \alpha_1 \left(\frac{m}{r^2} \right)^{2/3+\varepsilon} \sum_{j=1}^{c} n_j^{2/3} + \alpha_2 \left(m' + \sum_{j=1}^{c} n_j \right). \tag{7.4}$$

The above bound of $O(r)$ on the number of intersection points between a line $\ell \in \mathcal{L} \setminus \mathcal{L}_0$ and $\mathbf{V}(f)$ implies that each line enters $O(r)$ cells (a line has to intersect

---

[3]By $a \ll b$ we mean that the constant b is larger than some expression depending on the constant a. For example, we might require that $b = \Omega(a^a)$.

$\mathbf{V}(f)$ when moving from one cell to another). This implies $\sum_{j=1}^{c} n_j = O(nr)$. Combining this with Hölder's inequality (see the "Notation and inequalities" part of the introduction) yields

$$\sum_{j=1}^{c} n_j^{2/3} \leq \left(\sum_{j=1}^{c} n_j\right)^{2/3} \left(\sum_{j=1}^{c} 1\right)^{1/3} = O\left((nr)^{2/3} \cdot r^{2/3}\right) = O\left(n^{2/3} r^{4/3}\right). \qquad (7.5)$$

By combining (7.4) and (7.5), we obtain

$$\sum_{j=1}^{c} I(\mathcal{P}_j, \mathcal{L}_j) = O\left(\frac{\alpha_1 m^{2/3+\varepsilon} n^{2/3}}{r^{2\varepsilon}} + \alpha_2 nr\right) + \alpha_2 m'.$$

Combining this with (7.2) and (7.3) yields

$$I(\mathcal{P}, \mathcal{L}) = O\left(\frac{\alpha_1 m^{2/3+\varepsilon} n^{2/3}}{r^{2\varepsilon}} + \alpha_2 nr + m_0 r\right) + \alpha_2 m'.$$

By taking α_2 to be sufficiently large with respect to r and the constant in the $O(\cdot)$-notation, we get

$$I(\mathcal{P}, \mathcal{L}) = O\left(\frac{\alpha_1 m^{2/3+\varepsilon} n^{2/3}}{r^{2\varepsilon}} + \alpha_2 nr\right) + \alpha_2(m' + m_0)$$

$$= O\left(\frac{\alpha_1 m^{2/3+\varepsilon} n^{2/3}}{r^{2\varepsilon}} + \alpha_2 nr\right) + \alpha_2 m. \qquad (7.6)$$

From (7.1) we have $n = n^{2/3} n^{1/3} = O(m^{2/3} n^{2/3})$. By taking α_1 to be sufficiently large with respect to α_2, r, and the constant in the $O(\cdot)$-notation in (7.6), we obtain $O(\alpha_2 nr) \leq \frac{\alpha_1}{2} m^{2/3} n^{2/3}$. Similarly, by taking r to be sufficiently large with respect to ε and the constant in the $O(\cdot)$-notation in (7.6), we may assume that

$$O\left(\frac{\alpha_1 m^{2/3+\varepsilon} n^{2/3}}{r^{2\varepsilon}}\right) \leq \frac{\alpha_1}{2} m^{2/3+\varepsilon} n^{2/3}.$$

Combining this with (7.6) completes the induction step, and thus the proof of the theorem. $\square$

Remarks. (i) Without the extra ε in the exponent of the bound of Theorem 7.3, the induction step would have failed. Specifically, when using the induction hypothesis to sum up the incidences inside of the cells, we would have obtained an expression

that has the correct asymptotic value but with a leading constant that is larger than the one we started with. A similar situation always occurs when using constant-degree polynomial partitioning, and in some sense this is the main disadvantage of the technique.

(ii) Even though we do not apply the weak incidence bound (Lemma 7.1) in every cell as before, the proof still relies on this bound in a different place.

7.4   The Szemerédi–Trotter theorem in $\mathbb{C}^2$

Now that we have a basic understanding of constant-degree polynomial partitioning, we use this technique to handle a more difficult problem — the Szemerédi–Trotter theorem in $\mathbb{C}^2$. That is, we have a point set $\mathcal{P}$ and a set of lines $\mathcal{L}$, both in $\mathbb{C}^2$. One can think of a complex line as the zero set (over the complex numbers) of a linear bivariate polynomial with coefficients in $\mathbb{C}$.

Theorem 7.4. *Let $\mathcal{P}$ be a set of m points and let $\mathcal{L}$ be a set of n lines, both in $\mathbb{C}^2$. Then for any $\varepsilon > 0$, we have $I(\mathcal{P}, \mathcal{L}) = O_\varepsilon(m^{2/3+\varepsilon}n^{2/3} + m + n)$.*

Proof. We move from $\mathbb{C}^2$ to $\mathbb{R}^4$ by using the map ϕ defined in Section 7.1. That is, a point $(x_1 + iy_1, x_2 + iy_2) \in \mathbb{C}^2$ is considered as the point $(x_1, y_1, x_2, y_2) \in \mathbb{R}^4$, and the set $\mathcal{P}$ becomes a set of m points in $\mathbb{R}^4$. Given a line ℓ in $\mathbb{C}^2$, we can write $\ell = \mathbf{V}((a + ia')z_1 + (b + ib')z_2 + (c + ic'))$ for some constants $a, a', b, b', c, c' \in \mathbb{R}$. A point $(x_1 + iy_1, x_2 + iy_2) \in \mathbb{C}^2$ is in ℓ if and only if

$$(a + ia')(x_1 + iy_1) + (b + ib')(x_2 + iy_2) + (c + ic') = 0,$$

or equivalently,

$$bx_2 - b'y_2 + ax_1 - a'y_1 + c = 0 \quad \text{and} \quad b'x_2 + by_2 + a'x_1 + ay_1 + c' = 0.$$

Thus, when moving to $\mathbb{R}^4$ a line ℓ is defined by two linear equations. It is not difficult to verify that each equation defines a distinct hyperplane, and that these two hyperplanes are not parallel. That is, ℓ becomes a 2-flat in $\mathbb{R}^4$. We thus consider $\mathcal{L}$ as a set of n 2-flats in $\mathbb{R}^4$. Note that the set of incidences in $\mathcal{P} \times \mathcal{L}$ does not change when moving from $\mathbb{C}^2$ to $\mathbb{R}^4$. Since two lines in $\mathbb{C}^2$ intersect in at most one point, every two 2-flats of $\mathcal{L}$ intersect in at most one point. That is, the incidence graph of $\mathcal{P} \times \mathcal{L}$ contains no $K_{2,2}$. To complete the proof of the theorem, we derive an upper bound on the number of point-flat incidences in $\mathbb{R}^4$.

An incidence bound in $\mathbb{R}^4$**.** Imitating the proof of Theorem 7.3, we prove the point-flat incidence bound by induction on $m + n$. Specifically, we prove that for any fixed $\varepsilon > 0$ there exist constants α_1, α_2 such that

$$I(\mathcal{P}, \mathcal{L}) \leq \alpha_1 m^{2/3+\varepsilon} n^{2/3} + \alpha_2(m + n). \tag{7.7}$$

For the induction basis, the bound holds for small $m + n$ by choosing α_1 and α_2 sufficiently large. For the induction step, we recall that the incidence graph contains no copy of $K_{2,2}$. As before, the weak incidence bound (Lemma 7.1) implies $I(\mathcal{P}, \mathcal{L}) = O(m\sqrt{n} + n)$. When $m = O\left(\sqrt{n}\right)$ this implies $I(\mathcal{P}, \mathcal{L}) = O(n)$, which completes the proof. We may thus assume that

$$n = O\left(m^2\right). \tag{7.8}$$

We take r to be a sufficiently large constant, whose value depends on ε and will be determined below. Let f be an r-partitioning polynomial of $\mathcal{P}$. According to the polynomial partitioning theorem, f is of degree $O(r)$ and $\mathbf{V}(f)$ partitions $\mathbb{R}^4$ into connected cells, each containing at most m/r^4 points of $\mathcal{P}$. By Warren's theorem (Theorem 3.2), the number of cells is $c = O(r^4)$. As in the proof of Theorem 7.3, the relations between the constants of this proof are

$$2^\varepsilon \ll r \ll \alpha_2 \ll \alpha_1.$$

Denote the cells of the partition as $C_1, \ldots, C_c$. For $j = 1, \ldots, c$, put $\mathcal{P}_j = \mathcal{P} \cap C_j$ and let $\mathcal{L}_j$ denote the set of 2-flats of $\mathcal{L}$ that intersect C_j. Let $\mathcal{L}_0$ denote the subset of 2-flats of $\mathcal{L}$ that are contained in $\mathbf{V}(f)$, and let $\mathcal{P}_0 = \mathcal{P} \cap \mathbf{V}(f)$. Notice that

$$I(\mathcal{P}, \mathcal{L}) = I(\mathcal{P}_0, \mathcal{L}_0) + I(\mathcal{P}_0, \mathcal{L} \setminus \mathcal{L}_0) + \sum_{j=1}^{c} I(\mathcal{P}_j, \mathcal{L}_j).$$

Unlike in the planar case, we cannot rely on Bézout's theorem (Theorem 2.5) to bound the number of cells that are intersected by a 2-flat $h \in \mathcal{L}$. Instead, we notice that every cell of the partition that is intersected by a 2-flat $h \in \mathcal{L}$ corresponds to at least one connected component of $h \setminus \mathbf{V}(f)$. By Theorem 4.11 with $U = h$ and $W = \mathbf{V}(f)$, we get that h intersects $O(r^2)$ cells.

Bounding $\sum_{j=1}^{c} I(\mathcal{P}_j, \mathcal{L}_j)$**.** For $j = 1, \ldots, c$, put $m_j = |\mathcal{P}_j|$ and $n_j = |\mathcal{L}_j|$. We also set $m' = \sum_{j=1}^{c} m_j$, and recall that $m_j \leq m/r^4$ for every $1 \leq j \leq c$. By the induction hypothesis, we have

$$\sum_{j=1}^{c} I(\mathcal{P}_j, \mathcal{L}_j) \le \sum_{j=1}^{c} \left(\alpha_1 m_j^{2/3+\varepsilon} n_j^{2/3} + \alpha_2 (m_j + n_j) \right)$$

$$\le \alpha_1 \left(\frac{m}{r^4} \right)^{2/3+\varepsilon} \sum_{j=1}^{c} n_j^{2/3} + \alpha_2 \left(m' + \sum_{j=1}^{c} n_j \right). \qquad (7.9)$$

The above bound of $O(r^2)$ on the number of cells that are intersected by a 2-flat implies $\sum_{j=1}^{c} n_j = O(nr^2)$. Combining this with Hölder's inequality (see the "Notation and inequalities" part of the introduction) implies

$$\sum_{j=1}^{c} n_j^{2/3} = O\left(\left(nr^2 \right)^{2/3} \cdot r^{4/3} \right) = O\left(n^{2/3} r^{8/3} \right). \qquad (7.10)$$

By combining (7.9) and (7.10), we obtain

$$\sum_{j=1}^{c} I(\mathcal{P}_j, \mathcal{L}_j) \le O\left(\frac{\alpha_1 m^{2/3+\varepsilon} n^{2/3}}{r^{4\varepsilon}} + \alpha_2 nr^2 \right) + \alpha_2 m'.$$

Note that (7.8) implies $n = O(m^{2/3} n^{2/3})$. Thus, by taking α_1 to be sufficiently large with respect to α_2 and r, we have

$$\sum_{j=1}^{c} I(\mathcal{P}_j, \mathcal{L}_j) \le O\left(\frac{\alpha_1 m^{2/3+\varepsilon} n^{2/3}}{r^{4\varepsilon}} \right) + \alpha_2 m'.$$

Finally, by taking r to be sufficiently large with respect to ε and the constant of the $O(\cdot)$-notation, we have

$$\sum_{j=1}^{c} I(\mathcal{P}_j, \mathcal{L}_j) \le \frac{\alpha_1}{3} m^{2/3+\varepsilon} n^{2/3} + \alpha_2 m'. \qquad (7.11)$$

Bounding $I(\mathcal{P}_0, \mathcal{L} \setminus \mathcal{L}_0)$. If a 2-flat $h \in \mathcal{L}$ is not contained in $\mathbf{V}(f)$, then $\mathbf{V}(f) \cap h$ is at most one-dimensional. Specifically, $\mathbf{V}(f) \cap h$ is a variety of dimension at most one and of degree $O(r)$ (since it is defined by the linear equations of h and by f). We denote the set of these lower-dimensional varieties as $\Gamma = \{ h \cap \mathbf{V}(f) : h \in \mathcal{L} \setminus \mathcal{L}_0 \}$. We also set $|\mathcal{P}_0| = m_0$. Note that the incidence graph of $\Gamma \times \mathcal{P}_0$ contains no $K_{2,2}$. Thus, Lemma 7.2 implies

$$I(\mathcal{P}_0, \mathcal{L} \setminus \mathcal{L}_0) = I(\mathcal{P}_0, \Gamma) = O_r \left(m_0^{2/3} n^{2/3} + m_0 + n \right).$$

As before, by using (7.8) and taking α_1 and α_2 to be sufficiently large with respect to r and the constant of the $O(\cdot)$-notation, we have

$$I(\mathcal{P}_0, \mathcal{L} \setminus \mathcal{L}_0) \leq \frac{\alpha_1}{3} m^{2/3} n^{2/3} + \frac{\alpha_2}{2} m_0. \tag{7.12}$$

Bounding $I(\mathcal{P}_0, \mathcal{L}_0)$. Consider a point $p \in \mathcal{P}_0$ such that p is incident to two 2-flats $h, h' \in \mathcal{L}_0$. By performing a translation of $\mathbb{R}^4$ so that p becomes the origin, we can then think of h and h' as vector subspaces. Recall that h and h' intersect only in p, which implies that the vectors from these two spaces span all of $\mathbb{R}^4$. Since both h and h' are contained in $\mathbf{V}(f)$, their tangent planes at p (which are identical to h and h', respectively) are contained in the tangent hyperplane $T_p \mathbf{V}(f)$. Since two 2-flats cannot span $\mathbb{R}^4$ while also being contained in the same hyperplane, $T_p \mathbf{V}(f)$ is not well defined. We conclude that p is a singular point of $\mathbf{V}(f)$.

The above implies that at most one plane of $\mathcal{L}_0$ can be incident to a point of $\mathcal{P}_0$ that is a regular point of $\mathbf{V}(f)$. Such regular points contribute $O(m_0)$ incidences to $I(\mathcal{P}_0, \mathcal{L}_0)$.

Let V_{sing} be the set singular points of $\mathbf{V}(f)$. By Theorem 4.8, the set V_{sing} is a variety of dimension at most two and of degree $O_r(1)$. Thus, V_{sing} contains $O_r(1)$ 2-flats of $\mathcal{L}_0$ and these yield $O_r(m_0)$ incidences with the points of $\mathcal{P}_0$. The 2-flats of $\mathcal{L}_0$ that are not contained in V_{sing} intersect V_{sing} in varieties that are at most one-dimensional and of degree $O_r(1)$. By Lemma 7.2, the number of incidences that these lower-dimensional varieties contribute is $O_r\left(m^{2/3} n^{2/3} + m_0 + n\right)$. By combining the singular and regular cases, we get

$$I(\mathcal{P}_0, \mathcal{L}_0) = O_r\left(m^{2/3} n^{2/3} + m_0 + n\right).$$

Once again, by using (7.8) and taking α_1 and α_2 to be sufficiently large with respect to r and the constant of the $O(\cdot)$-notation, we have

$$I(\mathcal{P}_0, \mathcal{L}_0) \leq \frac{\alpha_1}{3} m^{2/3} n^{2/3} + \frac{\alpha_2}{2} m_0. \tag{7.13}$$

The induction step is obtained by combining (7.11), (7.12), and (7.13), and this in turn completes the proof of the theorem. $\qquad\square$

Zahl [108] removed the ε in the bound of Theorem 7.4 by using a more involved analysis. An alternative proof that does not rely on polynomial partitioning and leads to a bound with no ε was obtained by Tóth [101].

Solymosi and Tao proved an incidence result that is more general than the one stated for 2-flats in the proof of Theorem 7.4. The proof of the more general incidence is very similar to the one presented above.

Theorem 7.5 (Solymosi and Tao [90]). *Let $\mathcal{P}$ be a set of m points and let $\mathcal{V}$ be a set of n varieties of degree at most k, both in $\mathbb{R}^d$, such that:*

- *The dimension of every variety of $\mathcal{V}$ is at most $d/2$.*
- *The incidence graph contains no copy of $K_{s,t}$.*
- *There are no incidences between a point $p \in \mathcal{P}$ and a variety $U \in \mathcal{V}$ where p is a singular point of U.*
- *If a point $p \in \mathcal{P}$ is incident to two varieties $U, W \in \mathcal{V}$ and p is a regular point of both U and W, then the tangent spaces T_pU and T_pW intersect only in the origin.*

Then
$$I(\mathcal{P}, \mathcal{V}) = O_{\varepsilon,k,d,s,t}\left(m^{\frac{s}{2s-1}+\varepsilon} n^{\frac{2s-2}{2s-1}} + m + n\right).$$

Note that this result can be seen as a generalization of the planar incidence bound from Theorem 3.3.

7.5   (Optional) Arbitrary curves in $\mathbb{C}^2$

After using constant-degree polynomial partitioning to extend the Szemerédi–Trotter theorem to $\mathbb{C}^2$, we would like to derive a general point-curve incidence bound in $\mathbb{C}^2$. Unfortunately, the proof of Theorem 7.4 does not easily extend even to the case of circles. In that proof we relied on the property that if two 2-flats intersect in a point p, then their tangent planes at p also intersect in a single point. This property holds since if two lines in $\mathbb{C}^2$ intersect in a point p, then their tangent lines at p also intersect in a single point. Hardly any other family of curves in $\mathbb{C}^2$ has this property. For example, two circles in $\mathbb{C}^2$ may intersect in a point p while having the same tangent line at p.

Without the aforementioned tangent intersection property, we are no longer able to use the above argument for bounding the number of incidences with curves that are contained in the partition. To overcome this difficulty, we rely on a more analytic argument from [88]. We begin by defining some basic concepts from Differential Geometry. Since these concepts will not be used anywhere else in the book, we will be somewhat less rigorous than usual.

Intuitively, a set $M \subseteq \mathbb{R}^c$ is a *d-dimensional smooth manifold* if for every point $p \in M$ there exists an open set of M that contains p and "behaves" like a d-flat (specifically, the open set needs to be homeomorphic to an open subset of a d-flat). For our purpose, it suffices to state that any open subset of a d-dimensional variety V that does not contain any singular points of V is a d-dimensional smooth manifold.

For example, any open subset of a plane in $\mathbb{R}^3$ is a 2-dimensional smooth-manifold. An open subset of a circular conical surface is a smooth manifold if and only if it does not contain the apex (see Figure 4.2). For brevity, below we omit the word "smooth" and use the term manifold.

Let M be a d-dimensional manifold. The *tangent bundle* of M is a $2d$-dimensional manifold defined as

$$TM = \{(p, v) \in M \times (\mathbb{R}^d \setminus \{0\} \ : \ v \in T_pM\}.$$

For $d' < d$, let $E \subset TM$ be a $(d + d')$-dimensional sub-manifold of TM. We say that E is a d'-dimensional *sub-bundle* of TM if for every point $p \in M$ we have $(\{p\} \times T_pM) \cap E = \{p\} \times V$, where V is a d'-dimensional vector subspace of $T_pM = \mathbb{R}^d$. We denote the subspace V as $E(p) \subset \mathbb{R}^d$.

A *vector field* of a manifold $M \subset \mathbb{R}^c$ is an assignment of a vector of $\mathbb{R}^c$ to each point of M. That is, if X is a vector field on a manifold $M \subset \mathbb{R}^c$, then $X : M \to \mathbb{R}^d \setminus \{0\}$. As with any function, a vector field X is said to be *smooth* if X is infinitely differentiable. Note that a one-dimensional sub-bundle of a manifold M is a smooth vector field on M. Abusing notation, we sometimes write $X(p) = v$ instead of $X(p) = (p, v) \in TM$. If E is a sub-bundle of TM and $X: M \to TM$ is a vector field, we say that X *takes values in* E if $X(p) \in E$ for all $p \in M$.

We are only interested in *smooth* vector fields. As with manifolds, we will omit the word smooth for brevity. The following is a variant of the Picard–Lindelöf theorem (e.g., see [58]).

Theorem 7.6. *Let X be a vector field on a manifold M and let $p \in M$. Then for any sufficiently small $\varepsilon > 0$ there exists a unique smooth arc $\alpha : [-\varepsilon, \varepsilon] \to M$ starting at p whose tangent vectors are in X; that is, a unique arc α that solves the initial value problem*

$$\alpha(0) = p, \qquad \alpha'(t) = X\big(\alpha(t)\big) \quad \text{for all } t \in [-\varepsilon, \varepsilon]. \tag{7.14}$$

We will also briefly work with *complex* varieties. The definition of a variety remains identical in this case: Given a set of polynomials $f_1, \ldots, f_k \in \mathbb{C}[x_1, \ldots, x_d]$, the variety $\mathbf{V}_{\mathbb{C}}(f_1, \ldots, f_k)$ is defined as

$$\mathbf{V}_{\mathbb{C}}(f_1, \ldots, f_k) = \{(a_1, \ldots, a_d) \in \mathbb{C}^d : f_j(a_1, \ldots, a_d) = 0 \text{ for all } 1 \leq j \leq k\}.$$

To prevent confusion regarding what field a variety is in, we will sometimes also use the notation $\mathbf{V}_{\mathbb{R}}(f_1, \ldots, f_k)$ to refer to a real variety.

Irreducibility, dimension, and singular points of a complex varieties are defined in the same way as for real varieties. For example, a singular point of a variety U is a point where the rank of the Jacobian matrix of U is smaller than the dimension of U (see Section 4.3 for the full definition). For a variety U, recall that we denote the set of singular points of U as U_{sing}, and the set of regular points of U as U_{reg}. Bézout's theorem (Theorem 2.5) and Theorem 4.8 also remain valid for complex varieties.

In the lemma below we have a polynomial $f \in \mathbb{R}[x_1, y_1, x_2, y_2]$ and consider the set $\mathbf{V}_{\mathbb{R}}(f)_{\text{reg}} \backslash \mathbf{V}_{\mathbb{C}}(f)_{\text{sing}}$. This set consists of the regular points of the variety that is defined by f in $\mathbb{R}^4$ that are not singular points of the variety that is defined by f in $\mathbb{C}^4$. For example, let $f = x_1^2 + x_2^2 \in \mathbb{R}[x_1, x_2, x_3]$. Since $\mathbf{V}_{\mathbb{R}}(f)$ is the x_3-axis, the set $\mathbf{V}_{\mathbb{R}}(f)_{\text{reg}}$ is also the x_3-axis. Since $\mathbf{V}_{\mathbb{C}}(f)$ is a union of two planes that intersect in the x_3-axis, the set $\mathbf{V}_{\mathbb{C}}(f)_{sing}$ is also the x_3-axis. We conclude that $\mathbf{V}_{\mathbb{R}}(f)_{\text{reg}} \backslash \mathbf{V}_{\mathbb{C}}(f)_{\text{sing}} = \emptyset$.

Recall that the map $\phi : \mathbb{C}^2 \to \mathbb{R}^4$ is defined by $\phi(x_1 + iy_1, x_2 + iy_2) = (x_1, y_1, x_2, y_2)$. The following lemma will be our main tool for bounding the number of incidences on a constant-sized partitioning polynomial.

Lemma 7.7. *Consider $f \in \mathbb{R}[x_1, y_1, x_2, y_2]$ such that the variety $\mathbf{V}_{\mathbb{R}}(f)$ is three-dimensional. Then for every $p \in \mathbf{V}_{\mathbb{R}}(f)_{reg} \backslash \mathbf{V}_{\mathbb{C}}(f)_{sing}$, there is at most one irreducible complex curve $\gamma \subset \mathbb{C}^2$ with $p \in \phi(\gamma)_{reg}$ and $\phi(\gamma) \subset \mathbf{V}_{\mathbb{R}}(f)$.*

Proof. We set $M = \mathbf{V}_{\mathbb{R}}(f)_{\text{reg}} \backslash \mathbf{V}_{\mathbb{C}}(f)_{\text{sing}}$. By Lemma 4.8, the set $\mathbf{V}_{\mathbb{R}}(f)_{\text{reg}}$ is a three-dimensional manifold and the set $\mathbf{V}_{\mathbb{C}}(f)_{\text{sing}}$ is a variety of dimension at most two. Thus, the set M is a three-dimensional manifold in $\mathbb{R}^4$. The isomorphism ϕ turns multiplication by i in $\mathbb{C}^2$ into the linear transformation

$$J : \mathbb{R}^4 \to \mathbb{R}^4, \qquad J(x_1, y_1, x_2, y_2) = (-y_1, x_1, -y_2, x_2).$$

Notice that for every $u \in \mathbb{R}^4$ we have $J(J(u)) = -u$. Thus, for any linear subspace $U \subset \mathbb{R}^4$ we have $J(J(U)) = U$. Since J corresponds to multiplication by i in $\mathbb{C}^2$, a linear subspace $U \subseteq \mathbb{R}^4$ is J-invariant if and only if $U = \phi(U')$ for some complex subspace $U' \subseteq \mathbb{C}^2$. In particular, all J-invariant subspaces are even dimensional.

For every point $p \in M$ we define the linear subspace $E_p = T_p M \cap J^{-1}(T_p M)$. Intuitively, E_p is the largest linear subspace of $T_p M$ that is invariant under J. Since the linear subspace $T_p M$ is three-dimensional, it cannot be J-invariant. This implies that $J^{-1}(T_p M)$ is a different three-dimensional subspace, and thus E_p is a two-dimensional linear subspace. As p varies, the union of the $p \times E_p$ forms a two-dimensional sub-bundle E of the tangent bundle TM.

Fix a point $p \in M$, and let X be a vector field such that X is defined in an open neighbourhood $U \subseteq M$ of p, the restriction of X to U takes values in E, and

$X(p) \neq 0$. Note that if such a vector field does not exist, then no irreducible curve $\gamma \subset \mathbb{C}^2$ satisfies $p \in \phi(\gamma)_{\mathrm{reg}}$ and $\phi(\gamma) \subset \mathbf{V}_{\mathbb{R}}(f)$. We thus assume that X exists. By Theorem 7.6 there is a unique arc $\alpha : [-\varepsilon, \varepsilon] \to \gamma_{\mathrm{reg}}$ that solves (7.14) (for a sufficiently small $\varepsilon > 0$). Two distinct irreducible curves intersect in a zero-dimensional variety, and by Theorem 4.12 this variety consists of finitely many points. Thus, there could be at most one irreducible one-dimensional variety that contains γ. We denote this curve as α^* (if no such curve exists then we are done).

Consider an irreducible complex curve $\gamma \subset \mathbb{C}^2$ that satisfies $\phi(\gamma) \subset \mathbf{V}_{\mathbb{R}}(f)$ and a point $p \in \phi(\gamma)_{\mathrm{reg}}$. For a point $q \in \gamma$, note that if $\phi(q) \in \phi(\gamma)_{\mathrm{reg}}$ then the tangent space $T_q\gamma$ is a complex line in $\mathbb{C}^2$. Since $T_{\phi(q)}\phi(\gamma) = \phi(T_q\gamma)$, the tangent space $T_{\phi(q)}\phi(\gamma)$ is a J-invariant 2-flat in $\mathbb{R}^4$ that is contained in T_qM. This implies that $T_{\phi(q)}\phi(\gamma) = E_q$, so $X(q)$ is tangent to $\phi(\gamma)$. By applying Theorem 7.6 to the manifold $\phi(\gamma)_{\mathrm{reg}}$, and the restriction of X to $\phi(\gamma)_{\mathrm{reg}}$, we obtain an arc $\beta : [-\varepsilon', \varepsilon'] \to \phi(\gamma)_{\mathrm{reg}}$ that solves the same equation (7.14) (for a sufficiently small $0 < \varepsilon' \leq \varepsilon$). The uniqueness of the solution implies that α^* and β must have a common sub-arc around p.

Assume for contradiction that there exists another irreducible complex curve $\gamma' \subset \mathbb{C}^2$ that satisfies $\phi(\gamma') \subset \mathbf{V}_{\mathbb{R}}(f)$ and $p \in \phi(\gamma')_{\mathrm{reg}}$. By repeating the argument from the previous paragraph, we get another arc $\beta' : [-\varepsilon'', \varepsilon''] \to \phi(\gamma)_{\mathrm{reg}}$ such that α^* and β' have a common sub-arc α' around p. This in turn implies that β and β' have a common sub-arc around p. We got a contradiction to Bézout's theorem, which states that γ and γ' have a finite intersection. That is, at most one curve γ satisfies the above properties. $\square$

We are finally ready to study incidences with arbitrary curves in $\mathbb{C}^2$.

Theorem 7.8. *Let $\mathcal{P}$ be a set of m points and let Γ be a set of n distinct irreducible algebraic curves of degree at most k, both in $\mathbb{C}^2$. If the incidence graph of $\mathcal{P} \times \Gamma$ contains no copy of $K_{s,t}$, then for any $\varepsilon > 0$ we have*

$$I(\mathcal{P}, \Gamma) = O_{s,t,k,\varepsilon}\left(m^{\frac{s}{2s-1}+\varepsilon}n^{\frac{2s-2}{2s-1}} + m + n\right).$$

Proof. The first half of the proof is very similar to the proof of Theorem 7.3. A reader who wishes to avoid reading this repetition may like to skip to the part titled "Incidences on the partition".

From Lemma 7.1 we get the weak bound $I(\mathcal{P}, \Gamma) = O_{s,t}\left(mn^{1-\frac{1}{s}} + n\right)$. When $m = O(n^{1/s})$, this implies the bound $I(\mathcal{P}, \Gamma) = O(n)$. We may thus assume that

$$n = O\left(m^s\right). \tag{7.15}$$

We will prove by induction on $m + n$ that

$$I(\mathcal{P}, \Gamma) \leq \alpha_1 m^{\frac{s}{2s-1}+\varepsilon} n^{\frac{2s-2}{2s-1}} + \alpha_2(m + n),$$

where α_1, α_2 are sufficiently large constants. The base case where $m + n$ is small can be handled by choosing sufficiently large values of α_1 and α_2.

We set $\mathcal{P}^* = \{\phi(p) : p \in \mathcal{P}\}$ and $\Gamma^* = \{\phi(\gamma) : \gamma \in \Gamma\}$. Note that $\mathcal{P}^*$ is a set of m points in $\mathbb{R}^4$ and that Γ^* is a set of n two-dimensional varieties of degree $O_k(1)$ in $\mathbb{R}^4$. As in the proof of Theorem 7.3, we will bound $I(\mathcal{P}^*, \Gamma^*)$. Since $\phi : \mathbb{C}^2 \to \mathbb{R}^4$ is a bijection, we indeed have $I(\mathcal{P}, \Gamma) = I(\mathcal{P}^*, \Gamma^*)$.

Partitioning $\mathbb{R}^4$. Let f be an r-partitioning polynomial of $\mathcal{P}^*$, for a sufficiently large constant r. According to the polynomial partitioning theorem (Theorem 4.10), f is of degree $O(r)$ and $\mathbf{V}(f)$ partitions $\mathbb{R}^4$ into connected cells, each containing at most m/r^4 points of $\mathcal{P}^*$. By Warren's theorem (Theorem 3.2), the number of cells is $c = O(r^4)$. The asymptotic relations between the various constants in the proof are

$$2^{1/\varepsilon} \ll r \ll \alpha_2 \ll \alpha_1.$$

Let $C_1, \ldots, C_c$ be the cells of the partition. Let $\mathcal{V}_j$ be the set of varieties from Γ^* that intersect the interior of C_j and let $\mathcal{P}_j$ be the set of points $p \in \mathcal{P}^*$ such that $p^* \in C_j$. Let $m_j = |\mathcal{P}_j|$, $m' = \sum_{j=1}^{c} m_j$, and $n_j = |\mathcal{V}_j|$. By Theorem 4.10, $m_j = O(m/r^4)$ for every $1 \leq j \leq c$. Notice that

$$I(\mathcal{P}^*, \Gamma^*) = I(\mathcal{P}_0, \Gamma_0) + I(\mathcal{P}_0, \Gamma^* \setminus \Gamma_0) + \sum_{j=1}^{c} I(\mathcal{P}_j, \Gamma_j).$$

For every every $U \in \Gamma^*$, applying Theorem 4.11 with $W = \mathbf{V}(f)$ implies that U intersects $O_k(r^2)$ cells of $\mathbb{R}^4 \setminus \mathbf{V}(f)$. Therefore, $\sum_{j=1}^{c} n_j = O_k(nr^2)$. Combining this with Hölder's inequality implies

$$\sum_{j=1}^{c} n_j^{\frac{2s-2}{2s-1}} = O_k\left(\left(nr^2\right)^{\frac{2s-2}{2s-1}} c^{\frac{1}{2s-1}}\right) = O_k\left(n^{\frac{2s-2}{2s-1}} r^{\frac{4s}{2s-1}}\right).$$

By the induction hypothesis, we have

$$
\sum_{j=1}^{c} I(\mathcal{P}_j, \Gamma_j) \leq \sum_{j=1}^{c} \left(\alpha_1 m_j^{\frac{s}{2s-1}+\varepsilon} n_j^{\frac{2s-2}{2s-1}} + \alpha_2(m_j + n_j) \right)
$$

$$
\leq O_k \left(\alpha_1 m^{\frac{s}{2s-1}+\varepsilon} r^{-\frac{4s}{2s-1}-4\varepsilon} \sum_{j=1}^{c} n_j^{\frac{2s-2}{2s-1}} \right) + \sum_{j=1}^{c} \alpha_2(m_j + n_j)
$$

$$
\leq O_k \left(\alpha_1 r^{-\varepsilon} m^{\frac{s}{2s-1}+\varepsilon} n^{\frac{2s-2}{2s-1}} \right) + \alpha_2 \left(m' + O_k \left(nr^2 \right) \right).
$$

By (7.15), we have $n^{\frac{1}{2s-1}} = O\left(m^{\frac{s}{2s-1}} \right)$, which in turn implies $n = O\left(m^{\frac{s}{2s-1}} n^{\frac{2s-2}{2s-1}} \right)$. Thus, when α_1 is sufficiently large with respect to r and α_2, we have

$$
\sum_{j=1}^{c} I(\mathcal{P}_j, \Gamma_j) = O_k \left(\alpha_1 r^{-\varepsilon} m^{\frac{s}{2s-1}+\varepsilon} n^{\frac{s}{2s-1}} \right) + \alpha_2 m'.
$$

By taking r to be sufficiently large with respect to ε, k, and the implicit constant in the O-notation, we have

$$
\sum_{j=1}^{c} I(\mathcal{P}_j, \Gamma_j) \leq \frac{\alpha_1}{2} m^{\frac{s}{2s-1}+\varepsilon} n^{\frac{2s-2}{2s-1}} + \alpha_2 m'. \tag{7.16}
$$

Incidences on the partition. It remains to bound incidences with points that are on the partitioning hypersurface $\mathbf{V}(f)$. Let $\mathcal{P}_0 = \mathcal{P}^* \cap \mathbf{V}(f)$, let $m_0 = |\mathcal{P}_0| = m - m'$, and let $\Gamma_0 = \{ U \in \Gamma^* : U \subset \mathbf{V}(f) \}$. By Lemma 7.7, every point of $(\mathcal{P}_0 \cap \mathbf{V}_\mathbb{R}(f)_{\text{reg}}) \setminus \mathbf{V}_\mathbb{C}(f)_{\text{sing}}$ is a regular point of at most one variety of Γ_0. That is, these points form at most m_0 incidences with regular points of varieties of Γ_0. It remains to bound the number of incidences $(p, U) \in \mathcal{P}_0 \times \Gamma^*$ that satisfy one of the following:

- The point p is in $\mathbf{V}_\mathbb{R}(f)_{\text{sing}}$ or in $\mathbf{V}_\mathbb{C}(f)_{\text{sing}}$.
- The point p is a singular point of U.
- The variety U is not in Γ_0.

Let $S = \mathbf{V}_\mathbb{R}(f)_{\text{sing}} \cup (\mathbf{V}_\mathbb{C}(f)_{\text{sing}} \cap \mathbb{R}^4)$. We first consider the case where U is contained in S. Since S is a two-dimensional variety of degree $O_{k,r}(1)$, it contains $O_{k,r}(1)$ varieties of Γ^*. Thus, the total number of incidences with the varieties of Γ^* that are in S is $O_{k,r}(m_0)$.

Next, we set

$$\Gamma_0^* = \{U \cap S \,:\, U \in \Gamma_0 \text{ and } U \not\subseteq S\} \bigcup \{U \cap \mathbf{V}(f) \,:\, U \in \Gamma^* \backslash \Gamma_0\} \bigcup \{U_{\text{sing}} \,:\, U \in \Gamma^*\}.$$

Note that Γ_0^* is a set of $O(n)$ varieties in $\mathbb{R}^4$, each of dimension at most one and of degree $O_{r,k}(1)$. This is straightforward for the first two parts in the definition of Γ_0^*, and is implied by Theorem 4.8 for the third part. By Lemma 7.2 we obtain

$$I(\mathcal{P}_0, \Gamma_0^*) = O_{s,t,k,r}\left(m_0^{\frac{s}{2s-1}} n^{\frac{2s-2}{2s-1}} + m_0 + n\right).$$

Combining the three cases above gives

$$I(\mathcal{P}_0, \Gamma^*) = O_{s,t,k,r}\left(m_0^{\frac{s}{2s-1}} n^{\frac{2s-2}{2s-1}} + m_0 + n\right).$$

Taking α_1, α_2 to be sufficiently large with respect to s, t, k, r, and the constant of the O-notation, we have

$$I(\mathcal{P}_0, \Gamma^*) \le \frac{\alpha_1}{2} m^{k/(2k-1)} n^{(2k-2)/(2k-1)} + \alpha_2(m_0 + n).$$

Combining this with (7.16) completes the proof of the induction step. $\qquad \square$

7.6   Exercises

Problem 7.1. The following result is from [45].

> **Theorem.** *Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^2$. Then at most $n^2/6 - n/2 + 1$ lines contain three points of $\mathcal{P}$.*

Use this theorem to prove that the same result holds for any set of n points in $\mathbb{R}^d$ (for every integer $d \ge 3$).

7.7   Open problems

In this chapter we saw how constant-degree polynomial partitioning makes it easier to study incidence problems in dimension $d \ge 3$. In Chapter 11 we will see more advanced uses of this technique for incidences in $\mathbb{R}^d$. This technique has one obvious drawback — it adds an extra ε to the exponent of the incidence bound. As already mentioned, Zahl [108] removed the ε in the special case that arises in the proof of Theorem 7.4. In Section 11.5 we will mention some other cases in which this ε was removed.

Open Problem 7.1. Remove the ε from the exponent in the bound of Theorem 7.8. More generally, develop a technique for removing the extra ε from the exponent whenever constant-degree polynomial partitioning is used.

We saw in this chapter how to derive point-curve incidence bounds in the complex plane $\mathbb{C}^2$. Unfortunately, so far it is not known how to extend our techniques to higher dimensional complex spaces. Elekes and Szabó [33] derived a bound for arbitrary point-variety incidences in $\mathbb{C}^d$. Their analysis relies on older incidence tools, so their bounds are weaker than the bounds obtained in $\mathbb{R}^d$ using polynomial partitioning (see Chapter 11). Dvir and Gopi [28] and Zahl [109] studied a specific point-line incidence problem in $\mathbb{C}^d$, for any $d \geq 3$. However, at the moment we do not have a reasonable bound for what may be considered as the main point-line incidence problem in $\mathbb{C}^3$.

Open Problem 7.2. For every $d \geq 3$, extend the known point-variety incidence bounds in $\mathbb{R}^d$ to $\mathbb{C}^d$.

Chapter 8

The Elekes–Sharir–Guth–Katz Framework

> *"By the way, in case of something unexpected happens to me (car accident, plane crash, a brick on the top of my skull) I definitely ask you to publish anything we have, at your will." / György Elekes, in an email to Micha Sharir, a few years before he passed away.*

Elekes and Sharir used to think about the planar distinct distances problem. Around the turn of the millennium Elekes communicated to Sharir the basics of a reduction from this problem to a problem involving intersections of helices in $\mathbb{R}^3$. Later on, Elekes sent Sharir the above quote.

Elekes passed away in 2008 and, as he requested, Sharir then published their ideas. Before publishing, Sharir simplified the reduction so that it led to a problem involving intersections of parabolas in $\mathbb{R}^3$. Publishing the reduction, thereby exposing it for the first time to the general community, proved to be a good idea. Hardly any time had passed before Guth and Katz managed to apply it to almost completely solve the planar distinct distances problem. (Recall from Chapter 1 that the current best upper bound for the problem is $\Omega(n/\sqrt{\log n})$.)

Theorem 8.1 (Guth and Katz [51]). *Every set of n points in $\mathbb{R}^2$ determines $\Omega(n/\log n)$ distinct distances.*

To obtain this result, Guth and Katz further improved the reduction so that it led to a problem concerning intersections between lines in $\mathbb{R}^3$. We thus refer to this reduction as the *Elekes–Sharir–Guth–Katz framework* (or the *ESGK framework*, for short).

The original proof of Guth and Katz is rather involved. In the following two chapters we will present a proof of slightly weaker variant of Theorem 8.1. This variant was introduced by Guth [48], and it allows us to avoid several of the more technical parts of the original proof such as studying ruled surfaces and flat points.

Theorem 8.2. *For every $\varepsilon > 0$, any set of n points in $\mathbb{R}^2$ determines $\Omega(n^{1-\varepsilon})$ distinct distances.*

The purpose of the current chapter is to introduce the first step of the proof: the ESGK framework.

8.1 Warmup: Distances between points on two lines

Before presenting the ESGK framework, we begin with a different distinct distances problem that can be easily reduced to an incidence problem. While working on this simpler problem we will encounter a couple of ideas used in the ESGK framework.

In a *bipartite* distinct distances problem we have two point sets $\mathcal{P}_1$ and $\mathcal{P}_2$, and are interested in the number $D(\mathcal{P}_1, \mathcal{P}_2)$ of distinct distances between pairs from $\mathcal{P}_1 \times \mathcal{P}_2$. That is,

$$D(\mathcal{P}_1, \mathcal{P}_2) = \left| \{|pq| \ : \ p \in \mathcal{P}_1, q \in \mathcal{P}_2\} \right|$$

(where $|pq|$ denotes the distance between the points p and q).

We consider a planar bipartite problem where $\mathcal{P}_1$ is a set of m points that lie on a line ℓ_1 and $\mathcal{P}_2$ is a set of n points that lie on a different line ℓ_2. Without loss of generality, we assume that $n \geq m$. When the two lines are either parallel or orthogonal, the points can be arranged so that $D(\mathcal{P}_1, \mathcal{P}_2) = \Theta(n)$. Such constructions are illustrated in Figure 8.1.

On the other hand, when the two lines are neither parallel nor orthogonal, the current best construction for minimizing the number of distances yields $D(\mathcal{P}_1, \mathcal{P}_2) = \Theta\left(n^2 / \sqrt{\log n}\right)$ (see Elekes [31]). In this construction, we take ℓ_1 to be the x-axis and ℓ_2 to be the line $\mathbf{V}(y - x)$ (the line of slope one that is incident to the origin). We set

$$\mathcal{P}_1 = \{(j, 0) : 1 \leq j \leq n\} \quad \text{and} \quad \mathcal{P}_2 = \{(j, j) : 1 \leq j \leq m\}.$$

Recall that Theorem 1.8 states that the number of positive integers smaller than n that are the sum of two squares is $\Theta(n/\sqrt{\log n})$. In our case, every distance between a point of $\mathcal{P}_1$ and a point of $\mathcal{P}_2$ is of the form $\sqrt{d_x^2 + d_y^2}$, where both d_x and d_y are
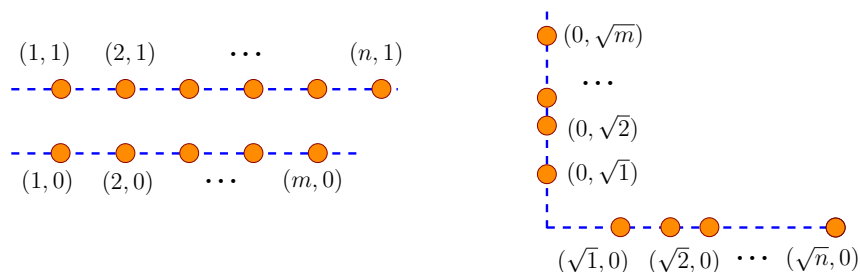
Figure 8.1: When the lines are either parallel or orthogonal, the points can be arranged so that $D(\mathcal{P}_1, \mathcal{P}_2) = \Theta(n)$.

integers between $-n$ and n. We may remove the square root since we only care about which distances are distinct, and not about the values of these distances. By Theorem 1.8 we obtain $D(\mathcal{P}_1, \mathcal{P}_2) = \Theta\left(n^2/\sqrt{\log n}\right)$.

We now derive the current best lower bound for the problem of distinct distances on two lines. Note the huge gap between the current best upper and lower bounds.

Theorem 8.3 ([87])**.** *Let ℓ_1 and ℓ_2 be lines in $\mathbb{R}^2$ that are neither parallel nor orthogonal. Let $\mathcal{P}_1$ be a set of n points on ℓ_1 and let $\mathcal{P}_2$ be a set of m points on ℓ_2. Then*

$$D(\mathcal{P}_1, \mathcal{P}_2) = \Omega\left(\min\left\{n^{2/3}m^{2/3}, n^2, m^2\right\}\right).$$

Proof. We begin by simplifying the problem. We rotate the plane so that ℓ_1 becomes the x-axis. We then translate the plane so that the intersection point $\ell_1 \cap \ell_2$ becomes the origin. Note that these transformations do not change the number of distinct distances. Let s denote the slope of ℓ_2 after the rotation. By the assumption, s is finite and nonzero. If the origin is in $\mathcal{P}_1$ or in $\mathcal{P}_2$ then we remove it from these sets. This can only decrease the number of distinct distances.

We set $D = D(\mathcal{P}_1, \mathcal{P}_2)$ and denote the D distinct distances in $\mathcal{P}_1 \times \mathcal{P}_2$ as $\delta_1, \ldots, \delta_D$. We also define the set of quadruples

$$Q = \{(a, p, b, q) \in \mathcal{P}_1 \times \mathcal{P}_2 \times \mathcal{P}_1 \times \mathcal{P}_2 \ : \ |ap| = |bq| > 0\}.$$

The quadruples of Q are ordered, so that (a, p, b, q) and (b, q, a, p) are considered as two distinct elements of Q. An example of a quadruple of Q is depicted in Figure 8.2. We prove the theorem by double counting $|Q|$. For every $1 \le j \le D$, let

$$E_j = \{(a, p) \in \mathcal{P}_1 \times \mathcal{P}_2 \ : \ |ap| = \delta_j\}.$$
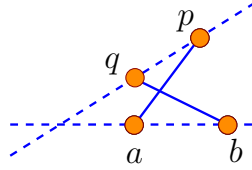
Figure 8.2: In a quadruple $(a, p, b, q) \in Q$, we have $|ap| = |bq| > 0$.

Since every pair of $\mathcal{P}_1 \times \mathcal{P}_2$ is in exactly one set E_j, we have

$$\sum_{j=1}^{D} |E_j| = |\mathcal{P}_1| \cdot |\mathcal{P}_2| = mn.$$

Note that the number of quadruples of Q that consist of two pairs of points at distance δ_j is $|E_j|^2$. This implies that $|Q| = \sum_{j=1}^{D} |E_j|^2$. By the Cauchy–Schwarz inequality

$$\left(\sum_{j=1}^{D} |E_j| \right)^2 \leq \left(\sum_{j=1}^{D} |E_j|^2 \right) \left(\sum_{j=1}^{D} 1 \right) = D \sum_{j=1}^{D} |E_j|^2.$$

By combining the above, we get

$$|Q| = \sum_{j=1}^{D} |E_j|^2 \geq \frac{1}{D} \left(\sum_{j=1}^{D} |E_j| \right)^2 = \frac{m^2 n^2}{D}. \tag{8.1}$$

Deriving an upper bound for $|Q|$. Consider a quadruple $(a, p, b, q) \in \mathcal{P}_1 \times \mathcal{P}_2 \times \mathcal{P}_1 \times \mathcal{P}_2$, and write $a = (a_x, 0), b = (b_x, 0), p = (p_x, sp_x)$, and $q = (q_x, sq_x)$. Recall that this quadruple is in Q if and only if $|ap| = |bq|$, or equivalently

$$(a_x - p_x)^2 + s^2 p_x^2 = (b_x - q_x)^2 + s^2 q_x^2. \tag{8.2}$$

For every pair $(p, q) \in \mathcal{P}_2^2$ we define a corresponding point $v_{pq} = (p_x, q_x) \in \mathbb{R}^2$. We define $\mathcal{P}' = \{v_{pq} : (p, q) \in \mathcal{P}_2^2\}$. Note that $\mathcal{P}'$ is a set of m^2 distinct points in $\mathbb{R}^2$. For every pair $(a, b) \in \mathcal{P}_1^2$ we define a corresponding hyperbola

$$\gamma_{ab} = \mathbf{V}\left(a_x^2 - b_x^2 - 2a_x x + 2b_x y + x^2(1 + s^2) - y^2(1 + s^2) \right).$$

Finally, we define $\mathcal{H} = \{\gamma_{ab} : (a, b) \in \mathcal{P}_1^2\}$. Note that $\mathcal{H}$ is a set of n^2 distinct hyperbolas.

For a quadruple (a, p, b, q), condition (8.2) is satisfied if and only if the point v_{pq} is incident to the hyperbola γ_{ab} (place the coordinates of v_{pq} in the polynomial defining γ_{ab}). Thus, to obtain an upper bound for $|Q|$ it suffices to obtain an upper bound for $I(\mathcal{P}', \mathcal{H})$. Recalling that $s \neq 0$, it is not difficult to verify that no hyperbola in $\mathcal{H}$ is degenerate (that is, the union of two lines).

By Bézout's theorem (Theorem 2.5), two non-degenerate hyperbolas intersect in at most four points. Thus, there is no $K_{5,2}$ in the incidence graph of $\mathcal{P}' \times \mathcal{H}$. This is a rather weak restriction, but we can improve it as follows. The roles of $(\ell_1, \mathcal{P})$ and $(\ell_2, \mathcal{P}_2)$ are symmetric and can be interchanged. That is, we can take pairs of $\mathcal{P}_1^2$ to form a point set $\overline{\mathcal{P}}'$ and pairs of $\mathcal{P}_2^2$ to form a set of hyperbolas $\overline{\mathcal{H}}$ (in the same way we created $\mathcal{P}'$ and $\mathcal{H}$). As before, Bézout's theorem implies that the incidence graph of $\overline{\mathcal{P}}' \times \overline{\mathcal{H}}$ contains no copy of $K_{5,2}$. The incidence graph of $\overline{\mathcal{P}}' \times \overline{\mathcal{H}}$ is identical to the incidence graph of $\mathcal{P}' \times \mathcal{H}$ but with the two sides of the graphs switched, since both describe the quadruples of Q. Thus, the incidence graph of $\mathcal{P}' \times \mathcal{H}$ contains no copy of $K_{2,5}$. We apply our point-curve incidence bound (Theorem 3.3) on $\mathcal{P}' \times \mathcal{H}$ with $s = 2$ and $t = 5$, to obtain

$$|Q| = O\left(|\mathcal{P}'|^{2/3}|\mathcal{H}|^{2/3} + |\mathcal{P}'| + |\mathcal{H}|\right) = O\left(m^{4/3}n^{4/3} + m^2 + n^2\right).$$

Combining this upper bound with the lower bound in (8.1) implies the assertion of the theorem. When comparing these two bounds, we split the analysis into three cases according to the term that dominates the upper bound for $|Q|$. Each case leads to a different term in the bound of the theorem. $\square$

8.2   The ESGK Framework

We are now ready to present the ESGK Framework. For this purpose, we return to the original problem of having a set $\mathcal{P}$ of n points in $\mathbb{R}^2$. Let x denote the number of nonzero distinct distances that are determined by pairs of points from $\mathcal{P}$. Similarly to the warmup problem, we consider the set

$$Q = \left\{(a, p, b, q) \in \mathcal{P}^4 \ : \ |ap| = |bq| > 0\right\}.$$

The quadruples in Q are ordered in the sense that (a, p, b, q), (b, q, a, p), (p, a, q, b), and the other possible permutations are all considered as distinct elements of Q. Note also that some of the four points in a quadruple may be identical. As before, the reduction is based on double counting $|Q|$, and we begin by deriving a lower bound.

We denote the set of nonzero distinct distances that are determined by $\mathcal{P} \times \mathcal{P}$ as $\delta_1, \ldots, \delta_x$. For every $1 \leq j \leq x$ we set

$$E_j = \left\{ (p, q) \in \mathcal{P}^2 \ : \ |pq| = \delta_j \right\}.$$

We consider (p, q) and (q, p) as two distinct pairs of E_j.

Since every ordered pair of distinct points of $\mathcal{P} \times \mathcal{P}$ is contained in a unique set E_j, we have $\sum_{j=1}^{x} |E_j| = n^2 - n$. Moreover, the number of quadruple of Q that consist of two pairs of points at distance δ_j is $|E_j|^2$. This implies that $|Q| = \sum_{j=1}^{x} |E_j|^2$. Applying the Cauchy-Schwarz inequality as in the previous section leads to

$$|Q| = \sum_{j=1}^{x} |E_j|^2 \geq \frac{1}{x} \left(\sum_{j=1}^{x} |E_j| \right)^2 = \frac{(n^2 - n)^2}{x} = \Theta \left(\frac{n^4}{x} \right). \qquad (8.3)$$

It remains to derive an upper bound for $|Q|$. Specifically, if we manage to derive the bound $|Q| = O(n^{3+\varepsilon})$, combining it with (8.3) would immediately imply $x = \Omega(n^{1-\varepsilon})$.

A transformation of $\mathbb{R}^2$ is a *rigid motion* if it preserves distances between points. The rigid motions of $\mathbb{R}^2$ are rotations, translations, reflections, and their combinations. A *proper rigid motion* is a rigid motion that also preserves orientation. That is, every ordered triple of points $(a, b, c) \in \mathbb{R}^2$ forms a left turn after applying the transformation if and only if it formed a left turn before the transformation. See Figure 8.3 for an example.
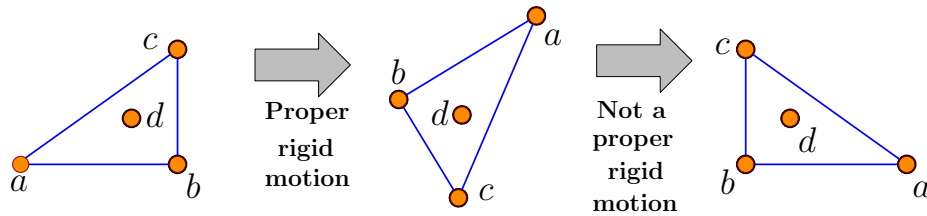
Figure 8.3: The second transformation is a rigid motion but not a proper one.

The only proper rigid motions of $\mathbb{R}^2$ are rotations and translations. Any combination of rotations and translations results in a single translation or in a single rotation (more details can be found in [94, Section 1.5] and in the exercises following it.)

For a pair of points $a, b \in \mathcal{P}$, consider the rotations that take a to b. The origin of such a rotation is equidistant from a and b. In other words, the centers of these rotations must all be on the perpendicular bisector of the segment ab. Conversely, every point on the perpendicular bisector of ab is the origin of a rotation that takes a to b. See Figure 8.4(a) for an illustration.

(a)   (b)   (c)
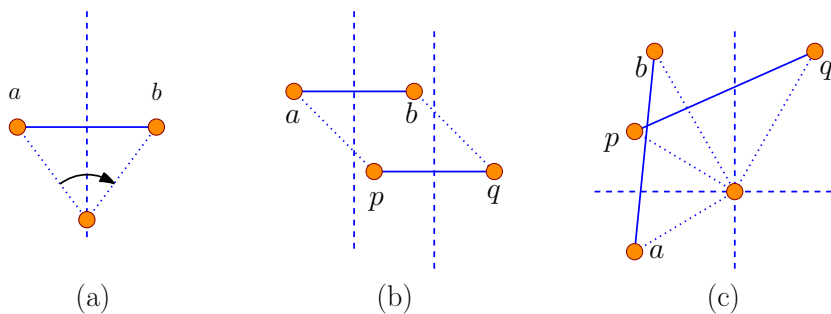
Figure 8.4: (a) The origin of the rotation is on the perpendicular bisector. (b) If the perpendicular bisectors are parallel then there is a translation taking ap to bq. (c) If the perpendicular bisectors intersect, there is a rotation taking ap to bq, and its origin is the intersection point.

Consider a quadruple $(a, p, b, q) \in Q$ and recall that by definition $|ap| = |bq| > 0$. We can always apply a translation that takes a to b and then rotate around the new position of a until p is taken to q. This translation followed by a rotation is a proper rigid motion taking ap to bq. To see that there is a unique proper rigid motion that takes ap to bq, we denote by ℓ_1 and ℓ_2 the perpendicular bisectors of the segments ab and pq, respectively. If ℓ_1 and ℓ_2 are parallel, then there is a unique translation taking ap to bq, and no rotations (for example, see Figure 8.4(b)). Similarly, if ℓ_1 and ℓ_2 intersect, then there is a unique rotation taking ap to bq, and no translations. The origin of this rotation is the point $\ell_1 \cap \ell_2$, and the angles of rotation from a to b and from p to q are equal because $|ap| = |bq|$ (see Figure 8.4(c)).

By the above, we have the following equivalent definition for Q: A quadruple (a, p, b, q) is in Q if and only if there exists a proper rigid motion τ that takes a to b and p to q. We say that the quadruple (a, p, b, q) *corresponds* to τ. To derive an upper bound for $|Q|$ it suffices to derive an upper bound for the number of quadruples from $\mathcal{P}^4$ that correspond to a proper rigid motion. In particular, we would like to prove that the number of such quadruples is $O(n^{3+\varepsilon})$. As already stated, combining this upper bound with (8.3) would lead to the required bound $x = \Omega(n^{1-\varepsilon})$.

We first bound the number of quadruples in Q that correspond to a translation. Given the first three points of a quadruple $(a, p, b, ?)$, there is at most one point in $\mathcal{P}$ that can complete it to a quadruple that corresponds to a translation. Thus, $O(n^3)$ quadruples in Q correspond to a translation.

Bounding the number of quadruples in Q that correspond to a rotation is significantly more difficult. A rotation of $\mathbb{R}^2$ can be described using three parameters: two coordinates for the center of rotation and another for the angle of rotation. Given a rotation with origin (o_x, o_y) and an angle of α, Guth and Katz [51] parameterized it

as $(o_x, o_y, \cot(\alpha/2)) \in \mathbb{R}^3$. With this parametrization, the set of rotations that take a fixed point $a \in \mathbb{R}^2$ to a fixed point $b \in \mathbb{R}^2$ form the following line in $\mathbb{R}^3$:

$$\ell_{ab} = \left\{ \left(\frac{a_x + b_x}{2}, \frac{a_y + b_y}{2}, 0 \right) + t \left(\frac{b_y - a_y}{2}, \frac{a_x - b_x}{2}, 1 \right) \ : \ t \in \mathbb{R} \right\}. \tag{8.4}$$

Showing that ℓ_{ab} is the set of rotations taking a to b is a standard technical calculation, so we postpone it to Section 8.3. A reader who is not interested in standard technical details might prefer skip Section 8.3.

The projection of ℓ_{ab} on the xy-plane is clearly a line, and by setting $t = 0$ in (8.4) we note that this projection contains the midpoint of a and b. Moreover, it is not difficult to verify that the projection is orthogonal to the line incident to both a and b. We conclude that the projection of ℓ_{ab} on the xy-plane is the perpendicular bisector of a and b (the set of points that are equidistant from a and b). That is, the line ℓ_{ab} is obtained by lifting the perpendicular bisector of a and b to a line in $\mathbb{R}^3$ incident to the midpoint of a and b and with a slope of $2/|ab|$ in the z-direction.

Consider a quadruple $(a, p, b, q) \in \mathcal{P}^4$ and the corresponding lines ℓ_{ab} and ℓ_{pq} in $\mathbb{R}^3$. If the intersection point $p = \ell_{ab} \cap \ell_{pq}$ exists, then p is the parametrization of a rotation that takes both a to b and p to q. That is, the quadruple (a, p, b, q) corresponds to a rotation (and is thus in Q) if and only if the lines ℓ_{ab} and ℓ_{pq} intersect. For an example, see Figure 8.5.
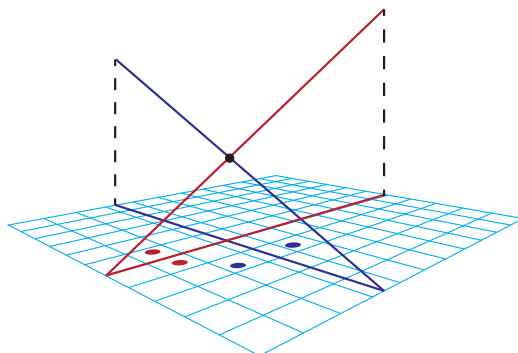
Figure 8.5: A quadruple of points in the plane, the two perpendicular bisectors, and their lifting to $\mathbb{R}^3$. Since the lifted lines intersect, there exists a rotation that takes one blue point to the other blue point and one red point to the other red point.

Let

$$\mathcal{L} = \{\ell_{ab} \ : \ (a, b) \in \mathcal{P}^2\}.$$

Note that $\mathcal{L}$ is a set of n^2 lines in $\mathbb{R}^3$,. By the above, there is a bijection between quadruples of Q that correspond to rotations and pairs of intersecting lines from $\mathcal{L}^2$.

Thus, an upper bound of $O(n^{3+\varepsilon})$ for the number of pairs of intersecting lines would imply $|Q| = O(n^{3+\varepsilon})$, as required. One can easily find sets of n^2 lines in $\mathbb{R}^3$ with $\Theta(n^4)$ intersecting pairs. However, as we will see in the following chapters, the lines of $\mathcal{L}$ have additional properties that prevent having too many intersecting pairs.

8.3   (Optional) Lines in the parametric space $\mathbb{R}^3$

In this section we study the parametrization of rotations of $\mathbb{R}^2$ that was presented in Section 8.2. That is, a rotation centered at a point $o \in \mathbb{R}^2$ and of angle α is parameterized as $\left(o_x, o_y, \cot \frac{\alpha}{2}\right) \in \mathbb{R}^3$. Given points $a, b \in \mathbb{R}^2$, we show that the set of parameterizations of rotations that take a to b is the line (8.4).

As stated above, the center of any rotation that takes a to b is on the perpendicular bisector of a and b. Let $c = ((a_x + b_x)/2, (a_y + b_y)/2)$ be the midpoint of the segment between a and b, and set $\delta = |ab| = \sqrt{(a_x - b_x)^2 + (a_y - b_y)^2}$. For example, see Figure 8.6.
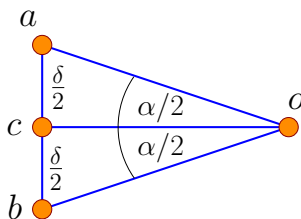
Figure 8.6: A rotation with origin o and angle α that takes a to b.

By definition, the perpendicular bisector of a and b is incident to the midpoint c and its slope is $s = (a_x - b_x)/(b_y - a_y)$. That is, the perpendicular bisector can be defined by

$$y - \frac{a_y + b_y}{2} = s\left(x - \frac{a_x + b_x}{2}\right).$$

Since o is incident to the perpendicular bisector of ab, we have

$$o_y - \frac{a_y + b_y}{2} = s\left(o_x - \frac{a_x + b_x}{2}\right). \tag{8.5}$$

We consider the case where $s \neq 0$, $o_x \geq c_x$, and $o_y \geq c_y$ (the other cases can be handled in a symmetric manner). We set $d_x = o_x - c_x$ and $d_y = o_y - c_y$. By (8.5) we

have $d_y = sd_x$, which implies

$$|co| = \sqrt{d_x^2 + d_y^2} = d_x\sqrt{1+s^2} = d_x\sqrt{\frac{(a_x - b_x)^2 + (b_y - a_y)^2}{(b_y - a_y)^2}} = \frac{\delta d_x}{b_y - a_y}. \qquad (8.6)$$

By looking at Figure 8.6, we notice that $|co| = \frac{\delta}{2}\cot\frac{\alpha}{2}$. Combining this with (8.6) gives

$$\frac{b_y - a_y}{2} \cdot \cot\frac{\alpha}{2} = d_x = o_x - c_x. \qquad (8.7)$$

By (8.5) and (8.7), we have that $\left(o_x, o_y, \cot\frac{\alpha}{2}\right)$ is on the line

$$\ell_{ab} = \left\{ \left(\frac{a_x + b_x}{2}, \frac{a_y + b_y}{2}, 0 \right) + t\left(\frac{b_y - a_y}{2}, \frac{a_x - b_x}{2}, 1 \right) \; : \; t \in \mathbb{R} \right\}.$$

Conversely, since we can choose o to be any point on the perpendicular bisector, any point on ℓ_{ab} is the parametrization of a rotation that takes a to b. Thus, ℓ_{ab} is exactly the set of parametrizations of the rotations that take a to b.

8.4 Exercises

Problem 8.1. Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^2$. We apply the ESKG framework to $\mathcal{P}$, obtaining a set of n^2 lines in $\mathbb{R}^3$. Let a, a', b, b', c, c' be distinct points of $\mathcal{P}$.
(a) Assume that every pair of the lines $\ell_{aa'}, \ell_{bb'}, \ell_{cc'}$ intersect. What does this say about the points a, a', b, b', c, c'?
(b) Assume that the three lines $\ell_{aa'}, \ell_{bb'}, \ell_{cc'}$ intersect at a single point. What new information do we have, beyond the properties obtained in part (a).

8.5 Open Problems

To obtain Theorem 8.1, Guth and Katz proved that every set of n points in $\mathbb{R}^2$ satisfies $|Q| = O(n^3 \log n)$ (where Q is defined as in Section 8.2). There are sets of n points for which this bound is tight, such as

$$\mathcal{P} = \{(a, b) \in \mathbb{Z}^2 \; : \; 1 \le a, b \le \sqrt{n}\}.$$

Thus, one cannot hope to eliminate the remaining gap between the current best upper and lower bounds for the distinct distances problem in $\mathbb{R}^2$ by deriving an improved upper bound for $|Q|$.

When considering the above set $\mathcal{P}$, we note that the gap between the lower and upper bounds for Q comes from the Cauchy-Schwarz argument in (8.3). More specifically, because of a non-tight application of Cauchy-Schwarz, (8.3) only leads to the bound $|Q| = \Omega(n^3\sqrt{\log n})$ (even though $|Q| = \Theta(n^3 \log n)$ in this case). We conclude that, to completely settle the distinct distances problem in $\mathbb{R}^2$, one has to change the ESKG framework.

Open Problem 8.1. Remove the $\sqrt{\log n}$ gap between the current best lower and upper bounds for the distinct distances problem in $\mathbb{R}^2$.

We now return to the problem of distinct distances between points on two lines, which was presented in Section 8.1. One reason for why this problem is important is that it has many generalizations (for example, see [72, 75, 76]). Improving the known bounds for this distances problem tends to lead to improvements for various generalizations. Quoting Hilbert [77]:

> *"The art of doing mathematics is finding that special case that contains all the germs of generality."*

Open Problem 8.2. Improve the current best bounds for the problem of distinct distances between points on two lines.

For example, consider the case where there are about n points on each line. In this case Theorem 8.3 implies that the number of distinct distances is $\Omega(n^{4/3})$, while it seems likely that the bound $\Omega(n^{2-\varepsilon})$ should also hold.

Chapter 9

Lines in $\mathbb{R}^3$

In Chapter 8 we reduced the planar distinct distances problem to a problem about pairs of intersecting lines in $\mathbb{R}^3$. In the current chapter we further reduce the problem to a point-line incidence problem in $\mathbb{R}^3$, and then solve this incidence problem. This completes the proof of the Guth-Katz distinct distances theorem. The incidence problem is where we use the simplified proof of Guth [48] instead of the original more involved argument. Due to this simplification, we prove Theorem 8.2 instead of the slightly stronger Theorem 8.1.

9.1 From line intersections to incidences

We begin by recalling where we stand in the proof of Theorem 8.2. For points $a, b \in \mathbb{R}^2$, we define the line $\ell_{a,b} \subset \mathbb{R}^3$ as

$$\ell_{ab} = \left(\frac{a_x + b_x}{2}, \frac{a_y + b_y}{2}, 0 \right) + t \left(\frac{b_y - a_y}{2}, \frac{a_x - b_x}{2}, 1 \right), \quad \text{for } t \in \mathbb{R}. \tag{9.1}$$

The line ℓ_{ab} is the set of parameterizations of the rotations of $\mathbb{R}^2$ that take a to b.

In the distinct distances problem, we are given a set $\mathcal{Q}$ of n points in $\mathbb{R}^2$ and wish to prove that $\mathcal{Q}$ spans many distinct distances (we reserve our standard notation $\mathcal{P}$ for a point set we will study in $\mathbb{R}^3$). For this purpose, we consider the set of n^2 lines

$$\mathcal{L} = \{\ell_{ab} : a, b \in \mathcal{Q}\}. \tag{9.2}$$

It is impossible for two lines of $\mathcal{L}$ to be identical. Indeed, if $\ell_{ab} = \ell_{cd}$ then every point on this line corresponds to a rotation of $\mathbb{R}^2$ taking both a to b and c to d.

However, unless $(a, b) = (c, d)$, there exists at most one rotation with such a property. Thus, any two lines of $\mathcal{L}$ can intersect in at most one point.

As we saw in Chapter 8, to prove that $\mathcal{Q}$ determines $\Omega(n^{1-\varepsilon})$ distinct distances, it suffices to prove that the number of pairs of intersecting lines in $\mathcal{L}$ is $O(n^{3+\varepsilon})$. Denote the number of such pairs as $X_\mathcal{L}$. We associate every pair of intersecting lines with their point of intersection. To derive an upper bound for $X_\mathcal{L}$, it suffices to go over every point of $\mathbb{R}^3$ and check how many pairs are associated with it. A point in $\mathbb{R}^3$ that is incident to r lines of $\mathcal{L}$ corresponds to $2\binom{r}{2}$ pairs of intersecting lines. We say that a point is r-*rich* if it is incident to at least r lines of $\mathcal{L}$ (as already defined in Section 1.2). For a positive integer r, let $\mathcal{P}_r(\mathcal{L})$ be the set of r-rich points. We perform a dyadic decomposition of $\mathcal{L}$, where for each j we consider points that are 2^j-rich but not 2^{j+1}-rich. For a fixed j, these points are contained in $\mathcal{P}_{2^j}(\mathcal{L})$, so the decomposition gives

$$X_\mathcal{L} < \sum_{j=1}^{2 \log n} |\mathcal{P}_{2^j}(\mathcal{L})| \cdot 2^{2j+2}. \tag{9.3}$$

To obtain the desired bound for $X_\mathcal{L}$, we will prove the following theorem.

Theorem 9.1. *Let $\mathcal{Q}$ be a set of n points in $\mathbb{R}^2$, and let $\mathcal{L}$ be the set of lines defined in (9.2). Then for any $\varepsilon > 0$ and $2 \le r \le n^2$, we have*

$$|\mathcal{P}_r(\mathcal{L})| = O\left(\frac{n^{3+\varepsilon}}{r^2}\right).$$

Let $0 < \varepsilon' < \varepsilon$. Applying Theorem 9.1 with ε' and combining the result with (9.3) gives

$$X_\mathcal{L} = O\left(\sum_{j=1}^{2 \log n} \frac{n^{3+\varepsilon'}}{2^{2j}} \cdot 2^{2j}\right) = O(n^{3+\varepsilon'} \log n) = O(n^{3+\varepsilon}).$$

That is, to prove Theorem 8.2 it remains to prove Theorem 9.1. We thus reduced the problem of bounding pairs of intersecting lines to a point-line incidence problem in $\mathbb{R}^3$.

Theorem 9.1 is false for arbitrary sets of n^2 lines in $\mathbb{R}^3$. For example, let h be an arbitrary plane in $\mathbb{R}^3$ and let $\mathcal{L}$ consist of n^2 generic lines in h. In this case every pair of lines intersect, so we cannot hope to prove that the number of intersecting pairs is $O(n^{3+\varepsilon})$. A similar issue occurs when $\mathcal{L}$ consists of n^2 lines in $\mathbb{R}^3$, all incident to the origin. Fortunately, since our set of lines $\mathcal{L}$ is constructed according to (9.2), it has additional properties.

Lemma 9.2. *Let $\mathcal{Q}$ be a set of n points in $\mathbb{R}^2$, and let $\mathcal{L}$ be the set of lines defined in* (9.2). *Then*
(i) Every point of $\mathbb{R}^3$ is incident to at most n lines of $\mathcal{L}$.
(ii) Every plane in $\mathbb{R}^3$ contains at most n lines of $\mathcal{L}$.

Proof. For a point $a \in \mathcal{Q}$, we set

$$\mathcal{L}_a = \{\ell_{ab} : b \in \mathcal{Q}\}.$$

(i) Consider two lines $\ell_{ab}, \ell_{ac} \in \mathcal{L}_a$. Recall from Chapter 8 that every point of ℓ_{ab} parameterizes a rotation of $\mathbb{R}^2$ that takes a to b, and similarly for the points of ℓ_{ac}. Since no rotation of $\mathbb{R}^2$ can take a to both b and c, the lines ℓ_{ab} and ℓ_{ac} do not intersect. Consider a point $p \in \mathbb{R}^3$. Since the lines of $\mathcal{L}_a$ do not intersect, at most one line of $\mathcal{L}_a$ is incident to p. Since $\mathcal{L} = \cup_{a \in \mathcal{Q}} \mathcal{L}_a$ and p is incident to at most one line from each family $\mathcal{L}_a$, we conclude that p is incident to at most n lines of $\mathcal{L}$.

(ii) When two lines are in the same plane in $\mathbb{R}^3$, they either intersect or are parallel. From part (i) of the current proof , we know that two lines from the same family $\mathcal{L}_a$ cannot intersect. By (9.1), the direction of the line ℓ_{ab} is $\left(\frac{b_y - a_y}{2}, \frac{a_x - b_x}{2}, 1\right)$, so two lines from $\mathcal{L}_a$ cannot be parallel. This implies that two lines from the same family cannot lie on a common plane. We conclude that every plane in $\mathbb{R}^3$ contains at most n lines of $\mathcal{L}$. $\qquad\square$

By Lemma 9.2(ii), when $r > n$ we have $\mathcal{P}_r(\mathcal{L}) = \emptyset$. It remains to prove Theorem 9.1 for the case of $2 \le r \le n$. For that purpose, we prove the following two results. For a set of lines $\mathcal{L}$ and a two-dimensional variety S, both in $\mathbb{R}^3$, we define

$$\mathcal{L}_S = \{\ell \in \mathcal{L} \;:\; \ell \subset S\}.$$

Theorem 9.3. *For any $\varepsilon > 0$, there exist sufficiently large constants C and D that satisfy the following. Let $\mathcal{L}$ be a set of n lines in $\mathbb{R}^3$, let $2 \le r \le 2n^{1/2}$, and let $r' = \lceil 9r/10 \rceil$. Then there exists a set $\mathcal{V}$ of varieties in $\mathbb{R}^3$ such that*

- *Every variety of $\mathcal{V}$ is irreducible, of dimension two, and of degree at most D.*
- *Every variety of $\mathcal{V}$ contains at least $n^{1/2+\varepsilon}$ lines of $\mathcal{L}$.*
- *$|\mathcal{V}| \le 2n^{1/2-\varepsilon}$.*
- *$|\mathcal{P}_r(\mathcal{L}) \setminus \cup_{S \in \mathcal{V}} \mathcal{P}_{r'}(\mathcal{L}_S)| \le Cn^{3/2+\varepsilon}/r^2$.*

Lemma 9.4. *Let $\mathcal{Q}$ be a set of n points in $\mathbb{R}^2$, and let $\mathcal{L}$ be the set of lines defined in* (9.2). *Then any irreducible two-dimensional variety $U \subset \mathbb{R}^3$ of degree at most D contains less than $2D^2n$ lines of $\mathcal{L}$.*

Using Theorem 9.3 and Lemma 9.4, it is straightforward to prove Theorem 9.1.

Proof of Theorem 9.1. By Applying Theorem 9.3 with $\mathcal{L}$, we obtain a set $\mathcal{V}$ of varieties in $\mathbb{R}^3$ of degree at most D, each containing at least $n^{1+\varepsilon}$ lines of $\mathcal{L}$. (Theorem 9.3 states that there are at least $n^{1/2+\varepsilon}$ lines in a vareity, but the theorem is stated for n lines and we are applying it to a set of n^2 lines.) By applying Lemma 9.4, we get that no surface of degree D can contain $n^{1+\varepsilon}$ lines of $\mathcal{L}$, so $\mathcal{V} = \emptyset$. When $\mathcal{V} = \emptyset$, Theorem 9.3 implies

$$|\mathcal{P}_r(\mathcal{L})| = |\mathcal{P}_r(\mathcal{L}) \setminus \cup_{S \in \mathcal{V}} \mathcal{P}_{r'}(\mathcal{L}_S)| \le C n^{3+\varepsilon}/r^2,$$

as required. $\square$

In Section 9.2 we prove Theorem 9.3, and in Section 9.3 we prove Lemma 9.4. Since the proof of Lemma 9.4 is rather technical, we mark Section 9.3 as optional. In particular, in this section we rely on some basic Differential Topology, as already introduced in (the optional) Section 7.5.

It is not difficult to show that the bound of Theorem 9.1 is close to tight in some cases.

Claim 9.5. *There exists a set $\mathcal{L}$ of n^2 lines in $\mathbb{R}^3$ such that $\mathcal{P}_3(\mathcal{L}) = \Theta(n^3)$ and every plane in $\mathbb{R}^3$ contains $O(n^{1/2})$ lines of $\mathcal{L}$.*

Proof. We imitate the proof of Claim 5.2. Let Π be a set of m generic planes in $\mathbb{R}^3$, for a parameter m that will be set below. By generic planes, we mean that no two planes are parallel, no three intersect in a line, and no four intersect in a point. Set

$$\mathcal{L} = \{h \cap h' \ : \ h, h' \in \Pi \text{ and } h \ne h'\}.$$

Since no three planes intersect in a line, $\mathcal{L}$ is a set of $\binom{m}{2}$ distinct lines. We may also assume that no two lines of $\mathcal{L}$ are parallel. We fix the value of m such that $|\mathcal{L}| = n^2$, and note that $m = \Theta(n)$. For any three distinct planes h, h', h'', the three lines $h \cap h'$, $h \cap h''$, and $h' \cap h''$ intersect at a distinct point. Thus, $|\mathcal{P}_3(\mathcal{L})| = \binom{m}{3} = \Theta(n^3)$. $\square$

9.2   Rich points in $\mathbb{R}^3$

The goal of this section is to prove Theorem 9.3. That is, in this section we study rich points of lines in $\mathbb{R}^3$. In Section 1.2 we briefly studied rich points of lines in $\mathbb{R}^2$. In particular, Lemma 1.15 states that the number of r-rich points in any set of n lines in $\mathbb{R}^2$ is $O\left(\frac{n^2}{r^3} + \frac{n}{r}\right)$. This bound easily extends to $\mathbb{R}^3$.

Claim 9.6. *Let* $\mathcal{L}$ *be a set of* n *lines in* $\mathbb{R}^3$ *and let* $r \geq 2$. *Then* $|\mathcal{P}_r(\mathcal{L})| = O\left(\frac{n^2}{r^3} + \frac{n}{r}\right)$.

Proof. Let h be a generic plane in $\mathbb{R}^3$, let $\pi : \mathbb{R}^3 \to h$ be the projection onto h, let $\mathcal{P}' = \{\pi(p) : p \in \mathcal{P}_r(\mathcal{L})\}$, and let $\mathcal{L}' = \{\pi(\ell) : \ell \in \mathcal{L}\}$. Since h is chosen generically, we may assume that every point of $\mathcal{P}_r(\mathcal{L})$ is projected onto a distinct point in h and that every line of $\mathcal{L}$ is projected onto a distinct line in h. In particular, $|\mathcal{L}'| = |\mathcal{L}|$ and $|\mathcal{P}'| = |\mathcal{P}_r(\mathcal{L})|$.

Applying Lemma 1.15 in h implies that $|\mathcal{P}_r(\mathcal{L})| = O\left(\frac{n^2}{r^3} + \frac{n}{r}\right)$. Noting that $\mathcal{P}' \subseteq \mathcal{P}_r(\mathcal{L}')$ leads to

$$|\mathcal{P}_r(\mathcal{L})| = |\mathcal{P}'| \leq |\mathcal{P}_r(\mathcal{L}')| = O\left(\frac{n^2}{r^3} + \frac{n}{r}\right).$$

$\square$

When r is large, the following argument gives a good upper bound for the number of r-rich points.

Lemma 9.7. *Let* $\mathcal{L}$ *be a set of* n *lines in* $\mathbb{R}^3$ *and let* $r > 2n^{1/2}$. *Then* $|\mathcal{P}_r(\mathcal{L})| \leq 2n/r$.

Proof. Set $M = |\mathcal{P}_r(\mathcal{L})|$ and write $\mathcal{P}_r(\mathcal{L}) = \{p_1, \ldots, p_M\}$. We assume for contradiction that $M \geq r/2$. By definition, at least r lines of $\mathcal{L}$ are incident to p_1. At least $r - 1$ lines of $\mathcal{L}$ are incident to p_2 but not to p_1. At least $r - 2$ lines of $\mathcal{L}$ are incident to p_3 but to neither of p_1 and p_2. Continuing in this manner, we obtain

$$n = |\mathcal{L}| \geq \sum_{j=0}^{r/2-1} (r - j) > (r/2) \cdot (r/2) = r^2/4. \tag{9.4}$$

Since this contradicts the assumption $r > 2n^{1/2}$, we get that $M < r/2$. Repeating the argument leading to (9.4) now gives

$$n = |\mathcal{L}| \geq \sum_{j=0}^{M-1} (r - j) > (r/2) \cdot M.$$

The assertion of the lemma is obtained by rearranging this inequality. $\square$

We will require the following variant of Problem 4.2(a).

Lemma 9.8. *Consider two distinct irreducible two-dimensional varieties* $U, W \subset \mathbb{R}^3$ *of respective degrees* d_U *and* d_W. *Then* $U \cap W$ *contains at most* $d_U \cdot d_W$ *lines.*

Proof. Since U and W are distinct and irreducible, $U \cap W$ is a variety of dimension at most one. We may assume that this intersection is of dimension exactly one, since otherwise it contains no lines. Since $U \cap W$ consists of a finite number of irreducible components, it contains a finite number of lines.

Let h be a generic plane in $\mathbb{R}^3$, such that h is neither U nor W, every line contained in $U \cap W$ intersects h in a distinct point, and h does not contain any of the components of $U \cap W$. Then $h \cap U$ is a variety of dimension at most one and degree at most $\deg U$, and $h \cap W$ is a variety of dimension at most one and degree at most $\deg W$. By applying Bezout's theorem (Theorem 2.5) inside of h, we obtain that $h \cap U \cap W$ consists of at most $d_U \cdot d_W$ points. Since every line in $U \cap W$ intersects h at a distinct point, there are at most $d_U \cdot d_W$ such lines. □

The following result is obtained by imitating the proof of Lemma 9.7. Problem 9.2 is about practicing the same technique in a graph theoretic context.

Lemma 9.9. *Let $\mathcal{L}$ be a set of n lines in $\mathbb{R}^3$. Let $\mathcal{V}$ be a set of irreducible varieties in $\mathbb{R}^3$ of dimension two and degree at most D, each containing at least X lines of $\mathcal{L}$. If $X > 2Dn^{1/2}$ then $|\mathcal{V}| \leq 2n/X$*

Proof. Set $M = |\mathcal{V}|$ and write $\mathcal{V} = \{S_1, \ldots, S_M\}$. We assume for contradiction that $M \geq X/2D^2$. By Lemma 9.8, for every distinct $S_j, S_k \in \mathcal{V}$, the intersection $S_j \cap S_k$ contains at most D^2 lines of $\mathcal{L}$.

By definition, at least X lines of $\mathcal{L}$ are contained in S_1. At least $X - D^2$ lines of $\mathcal{L}$ are contained in S_2 but not in S_1. At least $X - 2D^2$ lines of $\mathcal{L}$ are contained in S_3 but in neither S_1 nor S_2. Continuing in this manner, we obtain

$$n = |\mathcal{L}| \geq \sum_{j=0}^{X/2D^2-1} (X - j \cdot D^2) > (X/2D^2) \cdot (X/2) = X^2/4D^2. \qquad (9.5)$$

Since this contradicts the assumption $X > 2Dn^{1/2}$, we get that $M < X/2D^2$. Repeating the argument leading to (9.5) now gives

$$n = |\mathcal{L}| \geq \sum_{j=0}^{M-1} (X - j \cdot D^2) > (X/2) \cdot M.$$

The assertion of the lemma is obtained by rearranging this inequality. □

We are now ready to prove Theorem 9.3, and first repeat the statement of this theorem.

Theorem 9.3. *For any $\varepsilon > 0$, there exist sufficiently large constants C and D that satisfy the following. Let $\mathcal{L}$ be a set of n lines in $\mathbb{R}^3$, let $2 \le r \le 2n^{1/2}$, and let $r' = \lceil 9r/10 \rceil$. Then there exists a set $\mathcal{V}$ of varieties in $\mathbb{R}^3$ such that*

- *Every variety of $\mathcal{V}$ is irreducible, of dimension two, and of degree at most D.*
- *Every variety of $\mathcal{V}$ contains at least $n^{1/2+\varepsilon}$ lines of $\mathcal{L}$.*
- *$|\mathcal{V}| \le 2n^{1/2-\varepsilon}$.*
- *$|\mathcal{P}_r(\mathcal{L}) \setminus \cup_{S \in \mathcal{V}} \mathcal{P}_{r'}(\mathcal{L}_S)| \le Cn^{3/2+\varepsilon}/r^2$.*

Proof. The proof idea is similar to that of the constant-degree polynomial partitioning technique (as described in Chapter 7): We prove the theorem by induction on n, partition the space into a constant number of cells, and apply the induction hypothesis separately in each cell. Unlike the analysis in Chapter 7, there may exist cells in which the induction hypothesis cannot be applied. For example, the induction hypothesis does no apply in cells that do not satisfy the assumption $r \le 2n^{1/2}$. Addressing this issue requires some new ideas.

As already stated, the proof is by induction on n. For the induction basis, the claim holds for small n by taking C to be sufficiently large and $\mathcal{V} = \emptyset$. We move to prove the induction step, assuming that n is at least some large constant.

Set $m = |\mathcal{P}_r(\mathcal{L})|$. We apply the polynomial partitioning theorem (Theorem 3.1) with $\mathcal{P}_r(\mathcal{L})$ and a sufficiently large constant s. This yields a polynomial $f \in \mathbb{R}[x, y, z]$ of degree $O(s)$, such that each of connected component of $\mathbb{R}^3 \setminus \mathbf{V}(f)$ contains at most m/s^3 points of $\mathcal{P}_r(\mathcal{L})$. Denote the open cells of the partition as $C_1, \ldots, C_v$. By Theorem 3.2, we have that $v = O(s^3)$. For each $1 \le j \le v$, denote by $\mathcal{L}_j$ the set of lines of $\mathcal{L}$ that intersect C_j, and set $\mathcal{P}_j = \mathcal{P}_r(\mathcal{L}) \cap C_j$.

We set $n_j = |\mathcal{L}_j|$. For a line $\ell \subset \mathbb{R}^3$, we apply Bezout's theorem in a generic plane containing ℓ, obtaining that either $\ell \subset \mathbf{V}(f)$ or $|\mathbf{V}(f) \cap \ell| \le \deg f$. When travelling along ℓ, each point where we cross to a new cell of the partition is in $\mathbf{V}(f) \cap \ell$. This implies that every line of $\mathcal{L}$ intersects at most $\deg f + 1 = O(s)$ cells of $\mathbb{R}^3 \setminus \mathbf{V}(f)$. Therefore, $\sum_{j=1}^{v} n_j = O(ns)$.

For a parameter $0 < \alpha < 1$, we say that a cell C_j is α-*good* if $n_j \le \alpha n/s^2$. Note that the number of cells that are not α-good is $O(ns/(\alpha n/s^2)) = O(\alpha s^3)$. Since each cell contains at most m/s^3 points of $\mathcal{P}_r(\mathcal{L})$, the number of points that are in non-α-good cells is $O(\alpha m)$. We take α to be a sufficiently small constant so that at most $m/10$ points of $\mathcal{P}_r(\mathcal{L})$ are contained in non-α-good cells. Let $\mathcal{P}_B$ be the set of these points. For brevity, we say that a cell is *good* if it is α-good for the chosen value of α.

Rich points in good cells. Let G be the set of indices $1 \le j \le v$ for which C_j is a good cell. We would like to apply the induction hypothesis separately in each good

cell, but cells for which $r > 2n_j^{1/2}$ do not satisfy the assumption of the theorem. Let $G_r \subset G$ be the set of indices $j \in G$ that satisfy $r \leq 2n_j^{1/2}$. For every $j \in G_r$ we apply the induction hypothesis on $\mathcal{L}_j$, to obtain a set of varieties $\mathcal{U}_j$ such that

$$|\mathcal{U}_j| \leq 2n_j^{1/2-\varepsilon} \leq 2(\alpha n/s^2)^{1/2-\varepsilon}.$$

Since $C_j \cap \mathcal{P}_r(\mathcal{L}) \subset \mathcal{P}_r(\mathcal{L}_j)$, the induction hypothesis gives

$$\left|(C_j \cap \mathcal{P}_r(\mathcal{L})) \setminus \cup_{S \in \mathcal{U}_j} \mathcal{P}_{r'}(\mathcal{L}_S)\right| \leq \left|\mathcal{P}_r(\mathcal{L}_j) \setminus \cup_{S \in \mathcal{U}_j} \mathcal{P}_{r'}(\mathcal{L}_S)\right| \leq \frac{Cn_j^{3/2+\varepsilon}}{r^2}$$

$$\leq \frac{C(\alpha n/s^2)^{3/2+\varepsilon}}{r^2}. \quad (9.6)$$

For $j \in G \setminus G_r$ we have $r > 2n_j^{1/2}$, so we can apply Lemma 9.7 with $\mathcal{L}_j$. By also recalling that $r \leq 2n^{1/2}$, we obtain

$$|C_j \cap \mathcal{P}_r(\mathcal{L})| \leq |\mathcal{P}_r(\mathcal{L}_j)| \leq 2n_j/r < 2n/r \leq 4n^{3/2}/r^2.$$

By setting $\mathcal{U}_j = \emptyset$ and taking C to be sufficiently large with respect to α and s, we obtain (9.6) also in this case. That is, (9.6) holds for every $j \in G$. By summing up over every good cell, we obtain

$$\sum_{j \in G} \left|(C_j \cap \mathcal{P}_r(\mathcal{L})) \setminus \cup_{S \in \mathcal{U}_j} \mathcal{P}_{r'}(\mathcal{L}_S)\right| = O\left(\frac{s^3 \cdot C(\alpha n/s^2)^{3/2+\varepsilon}}{r^2}\right) = O\left(\frac{C(\alpha n)^{3/2+\varepsilon}}{s^{2\varepsilon} r^2}\right).$$

By taking s to be sufficiently large with respect to ε and to the constant of the $O(\cdot)$-notation, we get

$$\sum_{j \in G} \left|(C_j \cap \mathcal{P}_r(\mathcal{L})) \setminus \cup_{S \in \mathcal{U}_j} \mathcal{P}_{r'}(\mathcal{L}_S)\right| \leq \frac{Cn^{3/2+\varepsilon}}{20r^2}. \quad (9.7)$$

Rich points on the partition.   We denote the irreducible components of $\mathbf{V}(f)$ as $Z_1, Z_2, \ldots, Z_u$, and set $\mathcal{P}'_j = \mathbf{V}(f) \cap \mathcal{P}_r(\mathcal{L})$. Some points of $\mathcal{P}_r(\mathcal{L})$ may appear in several sets $\mathcal{P}'_j$, but this will not affect the analysis. Consider a component Z_j and a point $p \in \mathcal{P}'_j$ such that $p \notin \mathcal{P}_{r'}(\mathcal{L}_{Z_j})$. Since $r' = \lceil 9r/10 \rceil$, there are at least $r/10$ lines of $\mathcal{L}$ that are incident to p and are not contained in Z_j. By Bezout's theorem, every line of $\mathcal{L}$ that is not contained in Z_j intersects Z_j in at most $\deg Z_j$ points. Combining the above gives

$$\left|\mathcal{P}'_j \setminus \mathcal{P}_{r'}(\mathcal{L}_{Z_j})\right| \leq \frac{10n \cdot \deg Z_j}{r}.$$

Summing this over every Z_j, recalling that $r \leq 2n^{1/2}$, and taking C to be sufficiently large with respect to s, yields

$$\left| (\mathcal{P}_r(\mathcal{L}) \cap \mathbf{V}(f)) \setminus \bigcup_{j=1}^{u} \mathcal{P}_{r'}(\mathcal{L}_{Z_j}) \right| \leq \frac{10n \cdot \deg f}{r} \leq \frac{20n^{3/2} \cdot \deg f}{r^2} \leq \frac{Cn^{3/2}}{20r^2}. \qquad (9.8)$$

Let $\mathcal{V}'$ be a collection of varieties, consisting of every irreducible component of $\mathbf{V}(f)$, and of the varieties of $\mathcal{U}_j$ for every $j \in G$. By combining (9.7) and (9.8), and recalling that $\mathcal{P}_B$ is the set of points in the non-good cells, we have

$$\left| \mathcal{P}_r(\mathcal{L}) \setminus \bigcup_{S \in \mathcal{V}'} \mathcal{P}_{r'}(\mathcal{L}_S) \right| \leq |\mathcal{P}_B| + \frac{Cn^{3/2+\varepsilon}}{10r^2}. \qquad (9.9)$$

Taking D to be larger than $\deg f = O(s)$, we get that every variety in $\mathcal{V}'$ is of degree at most D. The number of varieties in $\mathcal{V}'$ is at most

$$O(s) + v \cdot 2(\alpha n/s^2)^{1/2-\varepsilon} = O(s^{2+2\varepsilon} n^{1/2-\varepsilon}).$$

Rich points in non-good cells. We almost obtained the statement of the theorem: The two remaining issues are the size of $|\mathcal{P}_B|$ and that some varieties of $\mathcal{V}'$ may contain fewer than $n^{1+\varepsilon}$ lines of $\mathcal{L}$. To resolve the first issue, we set $\mathcal{P}_1 = \mathcal{P}_r(\mathcal{L})$, $\mathcal{V}_1 = \mathcal{V}'$, and $\mathcal{P}_2 = \mathcal{P}_1 \setminus \cup_{S \in \mathcal{V}_1} \mathcal{P}_{r'}(\mathcal{L}_S)$. We then repeat the entire analysis for $\mathcal{P}_2$, and iteratively repeat this process. That is,

$$\mathcal{P}_j = \mathcal{P}_{j-1} \setminus \cup_{S \in \mathcal{V}_{j-1}} \mathcal{P}_{r'}(\mathcal{L}_S). \qquad (9.10)$$

By adapting (9.9) to the repeated argument and recalling that $|\mathcal{P}_B| \leq |\mathcal{P}_1|/10$, we obtain

$$|\mathcal{P}_j| \leq \frac{|\mathcal{P}_{j-1}|}{10} + \frac{Cn^{3/2+\varepsilon}}{10r^2}. \qquad (9.11)$$

Since every two lines intersect at most once, we have the trivial bound $|\mathcal{P}_1| \leq n^2$. Let $w = E \cdot \log n$ for some sufficiently large constant E. Then we have

$$|\mathcal{P}_w| < \frac{Cn^{3/2+\varepsilon}}{5r^2}.$$

Indeed, (9.11) implies that $|\mathcal{P}_{j+1}| < |\mathcal{P}_j|/2$ for every $|\mathcal{P}_j| \geq \frac{Cn^{3/2+\varepsilon}}{5r^2}$.

Set $\mathcal{V}^* = \bigcup_{j=1}^{w-1} \mathcal{V}_j$. Then

$$\left| \mathcal{P}_r(\mathcal{L}) \setminus \bigcup_{S \in \mathcal{V}^*} \mathcal{P}_{r'}(\mathcal{L}_S) \right| = |\mathcal{P}_w| < \frac{Cn^{3/2+\varepsilon}}{5r^2}. \tag{9.12}$$

For every $1 \le j \le w$ we have that $|\mathcal{V}_j| = O(s^{2+2\varepsilon}n^{1/2-\varepsilon})$. By summing over every j we obtain $|\mathcal{V}^*| = O(s^{2+2\varepsilon}n^{1/2-\varepsilon} \log n)$. We still have the issue that varieties of $\mathcal{V}^*$ may contain fewer than $n^{1+\varepsilon}$ lines of $\mathcal{L}$. Let $\mathcal{V}$ be the set of varieties of $\mathcal{V}^*$ that contain at least $n^{1/2+\varepsilon}$ lines of $\mathcal{L}$.

In the induction step we consider n larger than some sufficiently large constant, so we may assume that $n^\varepsilon > 2D$. We may thus apply Lemma 9.9 with $\mathcal{L}$ and $X = n^{1/2+\varepsilon}$, obtaining that

$$|\mathcal{V}| \le 2n/X = 2n^{1/2-\varepsilon}.$$

It remains to show that $\mathcal{P}_r(\mathcal{L}) \setminus \bigcup_{S \in \mathcal{V}} \mathcal{P}_{r'}(\mathcal{L}_S)$ is not too large. Recalling (9.12), it suffices to show that $\bigcup_{S \in \mathcal{V}^* \setminus \mathcal{V}} \mathcal{P}_{r'}(\mathcal{L}_S)$ is not too large. We perform a dyadic decomposition, defining $\mathcal{V}_j^*$ to be the set of varieties of $\mathcal{V}^*$ that contain at least 2^j lines of $\mathcal{L}$ and less than 2^{j+1} such lines. Note that $\mathcal{V}^* \setminus \mathcal{V} = \bigcup_{j=0}^{\log n^{1/2+\varepsilon}} \mathcal{V}_j^*$.

Consider a variety $S \in \mathcal{V}_j^*$. By Claim 9.6, the number of r'-rich points in a set of less than 2^{j+1} lines is

$$O\left(\frac{2^{2j}}{(r')^3} + \frac{2^j}{r'} \right) = O\left(\frac{2^{2j}}{r^3} + \frac{2^j}{r} \right). \tag{9.13}$$

When $2^j > 2Dn^{1/2}$, applying Lemma 9.9 with $\mathcal{L}$ and $X = 2^j$ implies that $|\mathcal{V}_j^*| \le 2n/2^j$. When $2^j \le 2Dn^{1/2}$, we use the trivial bound $|\mathcal{V}_j^*| \le |\mathcal{V}^*| = O(s^{2+2\varepsilon}n^{1/2-\varepsilon} \log n)$. Combining these two bounds with (9.13) implies

$$\sum_{S \in \mathcal{V}^* \setminus \mathcal{V}} |\mathcal{P}_{r'}(\mathcal{L}_S)| = \sum_{j=0}^{\log n^{1/2+\varepsilon}} \sum_{S \in \mathcal{V}_j^*} |\mathcal{P}_{r'}(\mathcal{L}_S)| = O\left(\sum_{j=0}^{\log n^{1/2+\varepsilon}} |\mathcal{V}_j^*| \cdot \left(\frac{2^{2j}}{r^3} + \frac{2^j}{r} \right) \right)$$

$$= O\left(\sum_{j=0}^{\log 2Dn^{1/2}} s^{2+2\varepsilon}n^{1/2-\varepsilon} \log n \cdot \left(\frac{2^{2j}}{r^3} + \frac{2^j}{r} \right) + \sum_{j=\log 2Dn^{1/2}}^{\log n^{1/2+\varepsilon}} \frac{2n}{2^j} \cdot \left(\frac{2^{2j}}{r^3} + \frac{2^j}{r} \right) \right)$$

$$= O\left(\left(\frac{s^{2+2\varepsilon}D^2n^{3/2-\varepsilon} \log n}{r^3} + \frac{s^{2+2\varepsilon}Dn^{1-\varepsilon} \log n}{r} \right) + \left(\frac{n^{3/2+\varepsilon}}{r^3} + \frac{\log n}{r} \right) \right).$$

Recall that in the induction hypothesis n is assumed to be larger than some sufficiently large constant, and that $r \le 2n^{1/2}$. By also taking C to be sufficiently

large with respect to the constant of the $O(\cdot)$-notation, s, and D, we obtain

$$\sum_{S \in \mathcal{V}^* \setminus \mathcal{V}} |\mathcal{P}_{r'}(\mathcal{L}_S)| \leq \frac{Cn^{3/2+\varepsilon}}{10r^2}.$$

Combining this with (9.12) leads to

$$\left| \mathcal{P}_r(\mathcal{L}) \setminus \bigcup_{S \in \mathcal{V}} \mathcal{P}_{r'}(\mathcal{L}_S) \right| < \frac{Cn^{3/2+\varepsilon}}{r^2}.$$

This completes the induction step, and the proof of the theorem. $\qquad \square$

9.3 (Optional) Lines in a two-dimensional surface

The purpose of this section is to prove Lemma 9.4. That is, we study the structure of lines that are contained in a two-dimensional variety in $\mathbb{R}^3$. Throughout the section we will use definitions and tools from Section 7.5, such a vector fields and tangent bundles. We will repeat the meaning of some of these definitions, but not all.

We begin with the following structural lemma. For a point $a \in \mathbb{R}^2$ we define $\overline{\mathcal{L}}_a = \{\ell_{a,b} : b \in \mathbb{R}^2\}$, where $\ell_{a,b}$ is defined as in (9.1). A *vector field* V of $\mathbb{R}^3$ is a map $V : \mathbb{R}^3 \to \mathbb{R}^3 \setminus \{0\}$. That is, V associates a vector with each point of $\mathbb{R}^3$.

Lemma 9.10. *For a point $a \in \mathbb{R}^2$:*
(i) Every point in $\mathbb{R}^3$ is incident to exactly one line of $\overline{\mathcal{L}}_a$.
(ii) There exists a vector field $V_a : \mathbb{R}^3 \to \mathbb{R}^3 \setminus \{0\}$ with the following properties: For every $x \in \mathbb{R}^3$ the vector $V_a(x)$ has the direction of the unique line of $\overline{\mathcal{L}}_a$ incident to x. Each of the three coordinates of $V_a(x)$ is a polynomial of degree at most 1 in the coordinates of a and of degree at most 2 in the coordinates of x.

Proof. (i) To distinguish between points in $\mathbb{R}^2$ and points in $\mathbb{R}^3$, we write $a = (a_x, a_y) \in \mathbb{R}^2$ but denote the coordinates of $x \in \mathbb{R}^3$ as x_1, x_2, x_3. For a point $b \in \mathbb{R}^2$, by inspecting (9.1) we note that ℓ_{ab} can be defined by

$$2x_1 = a_x + b_x + x_3(b_y - a_y),$$
$$2x_2 = a_y + b_y + x_3(a_x - b_x).$$

We rewrite this as a system of equations in the coordinates of b:

$$\begin{pmatrix} 1 & x_3 \\ -x_3 & 1 \end{pmatrix} \begin{pmatrix} b_x \\ b_y \end{pmatrix} = (2x_1 - a_x + a_y x_3, 2x_2 - a_y - a_x x_3).$$

The determinant of the matrix on the left side is $1+x_3^2 > 0$. Since this determinant is nonzero, for any fixed $x \in \mathbb{R}^3$ there is a unique b that solves the above system. That is, for every $x \in \mathbb{R}^3$ there is a unique line $\ell_{ab} \in \overline{\mathcal{L}}_a$ that is incident to x.

(ii) One can easily solve the above system for b by using Cramer's rule. This leads to

$$b_x = \frac{2x_1 - a_x + a_y x_3 - x_3(2x_2 - a_y - a_x x_3)}{x_3^2 + 1},$$

$$b_y = \frac{2x_2 - a_y - a_x x_3 + x_3(2x_1 - a_x + a_y x_3)}{x_3^2 + 1}.$$

By (9.1), the direction of $\ell_{a,b}$ is $\left(\frac{b_y - a_y}{2}, \frac{a_x - b_x}{2}, 1\right)$, or equivalently $(b_y - a_y, a_x - b_x, 2)$. Plugging in the above values for b_x and b_y gives

$$\left(\frac{2x_2 - a_y - a_x x_3 + x_3(2x_1 - a_x + a_y x_3)}{x_3^2 + 1} - a_y, \right.$$

$$\left. a_x - \frac{2x_1 - a_x + a_y x_3 - x_3(2x_2 - a_y - a_x x_3)}{x_3^2 + 1}, 2 \right).$$

Multiplying each coordinate by $x_3^2 + 1$ leads to

$$V_a(x) = \left(2x_2 - a_y - a_x x_3 + x_3(2x_1 - a_x + a_y x_3) - a_y(x_3^2 + 1), \right.$$

$$\left. a_x(x_3^2 + 1) - 2x_1 - a_x + a_y x_3 - x_3(2x_2 - a_y - a_x x_3), 2(x_3^2 + 1) \right).$$

We conclude that the vector field $V_a(x)$ satisfies all of the required properties. $\quad\square$

We next recall a few properties of varieties from Chapter 4. The *gradient* of a polynomial $f \in \mathbb{R}[x_1, x_2, x_3]$ is

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \right).$$

Let $U \subset \mathbb{R}^3$ be a two-dimensional variety and let $f \in \mathbb{R}[x_1, x_2, x_3]$ satisfy $\mathbf{I}(U) = \langle f \rangle$. Then a point $p \in U$ is a singular point of U if and only if $\nabla f(p) = (0, 0, 0)$. If p is a regular point of U then $\nabla f(p)$ is orthogonal to the tangent plane $T_p U$.

We are now ready to prove Lemma 9.4, and first repeat the statement of this lemma.

Lemma 9.4. *Let $\mathcal{Q}$ be a set of n points in $\mathbb{R}^2$, and let $\mathcal{L}$ be the set of lines defined in (9.2). Then any irreducible two-dimensional variety $U \subset \mathbb{R}^3$ of degree D contains less than $2D^2 n$ lines of $\mathcal{L}$.*

Proof. The case where U is a plane was already proved in Lemma 9.2. We thus assume that $D > 1$. To prove the lemma, we will show that at most one line family $\overline{\mathcal{L}}_a$ can have many of its lines contained in U.

Since U is an irreducible hypersurface of degree D, there exists $f \in \mathbb{R}[x_1, x_2, x_3]$ of degree D such that $\langle f \rangle = \mathbf{I}(U)$ (see Lemma 4.5). Consider a point $a \in \mathcal{Q}$, a line $\ell \in \overline{\mathcal{L}}_a$ contained in U, and the vector field V_a defined in Lemma 9.10(b). Let p be a regular point of U that is also incident to ℓ. Since $\ell \subset U$, this line is also contained in the tangent plane $T_p U$, which in turn implies that $V_a(p) \cdot \nabla f(p) = 0$. When p is a singular point we still have that $V_a(p) \cdot \nabla f(p) = 0$, since in this case $\nabla f(p) = (0, 0, 0)$. To recap, every line of $\overline{\mathcal{L}}_a$ that is contained in U is also contained in $\mathbf{V}(V_a \cdot \nabla f)$.

Assume that at least $2D^2$ lines of $\overline{\mathcal{L}}_a$ are contained in U, and set $W_a = \mathbf{V}(V_a \cdot \nabla f)$. By the previous paragraph, the intersection $W_a \cap U$ contains at least $2D^2$ lines. By Lemma 9.10(b), each part of V_a is of degree at most 2 in p. Since every part of ∇f is of degree at most $D - 1$, we get that $W_a \cap U$ is of degree at most $D + 1$. Since $\deg f = D$ and the intersection $W_a \cap U$ contains at least $2D^2 > D(D+1)$ lines, Lemma 9.8 implies that U and W_a have a common component. Since U is irreducible, we get that $U \subseteq W_a$.

We next assume that there exist distinct points $a, b \in \mathcal{Q}$ such that each of $\overline{\mathcal{L}}_a$ and $\overline{\mathcal{L}}_b$ has at least $2D^2$ lines in U. By Lemma 9.10(b), the polynomial $V_a \cdot \nabla f$ is linear in the coordinates of a (some coordinates of a may not appear in $V_a \cdot \nabla f$). Since both $V_a \cdot \nabla f$ and $V_b \cdot \nabla f$ vanish on U, so does $V_c \cdot \nabla f$ for any c in the affine span of a and b.

Let p be a regular point of U (recall from Theorem 4.8 that almost every point of U is regular), and let U_p be an open neighborhood of p in U that consists of regular points of U. Let c be in the affine span of a and b. By the preceding paragraph, $V_c \cdot \nabla f$ vanishes on U. This implies that the restriction of V_c to U_p is a sub-bundle of the tangent bundle TU_p. We denote this restriction of V_c as V_c^p. By the Picard–Lindelöf theorem (Theorem 7.6) with the manifold U_p and vector field V_c^p, there is a unique arc $\alpha : [-\varepsilon, \varepsilon] \to U_p$ that solves (7.14) (for a sufficiently small $\varepsilon > 0$). By the Picard–Lindelöf theorem with $\mathbb{R}^3$ and V_c, there exists a unique arc $\beta : [-\varepsilon', \varepsilon'] \to U_p$ that solves (7.14). Note that β clearly defines a segment of a line of $\overline{\mathcal{L}}_c$. Since V_c^p is a restriction of V_c, the curves defined by α and β coincide in some small neighborhood

of p. That is, in a sufficiently small neighborhood of p, the curve described by α is a segment of a line of $\overline{\mathcal{L}}_c$.

In the preceding paragraph we proved that U contains a segment of a line $\ell \in \overline{\mathcal{L}}_c$, and this segment is incident to p. Since U is a variety, it must contain ℓ (for example, by Bezout's theorem). This holds for every c in the affine span of a and b. Recalling that no line in $\mathbb{R}^3$ is contained in more than one family $\overline{\mathcal{L}}_c$, we get that there are infinitely many lines contained in U and incident to p. Since p is a regular point of U, all of these lines are contained in the plane T_pU. Applying Lemma 9.8 with U and T_pU implies that these two varieties have a common component. Since U is irreducible and T_pU is a plane, the two varieties are identical, which contradicts the assumption $D > 1$. Thus, there cannot exist distinct $a, b \in \mathcal{Q}$ such that each of $\overline{\mathcal{L}}_a$ and $\overline{\mathcal{L}}_b$ has at least $2D^2$ lines in U.

By the above, for at most one $a \in \mathcal{Q}$ the line family $\mathcal{L}_a$ has $2D^2$ or more lines in U. At most $n - 1$ lines from this family are in $\mathcal{L}$. For every other $b \in \mathcal{Q}$, at most $2D^2 - 1$ lines of $\mathcal{L}_b$ are in U. Summing up, we conclude that the number of lines of $\mathcal{L}$ that are in U is at most $(n - 1) + (n - 1)(2D^2 - 1) < 2nD^2$. $\square$

9.4   Exercises

Problem 9.1.
(i) Construct a set of n lines in $\mathbb{R}^2$ with $\Theta(n)$ points that are $n^{1/3}$-rich. (Hint: You might like to start with the construction of Claim 1.3 with $m = n$.)
(ii) Construct a set of n lines in $\mathbb{R}^3$ such that no plane contains more than $n^{1/2}$ of the lines and the number of $n^{1/6}$-rich points is $\Theta(n)$.

Problem 9.2. Consider a coloring of the edges of the complete graph K_n with the following properties: (i) No vertex is incident to two edges of the same color. (ii) There are no two colors c_1, c_2 and two vertices v, u such that both v and u are incident to an edge of color c_1 and to an edge of color c_2 (see Figure 9.1). Prove that the coloring consists of $\Theta(n^2)$ colors.
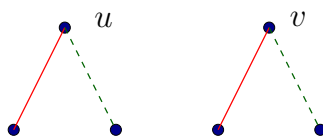

Figure 9.1: Both v and u are incident to an edge of color c_1 and to an edge of color c_2.

Problem 9.3. Consider a distinct distances variant for triangles, where two triangles are distinct if they are not congruent. We say that a point set *spans* a triangle if the three vertices of the triangle are points of the set. Prove that any set of n points in $\mathbb{R}^2$ spans $\Omega\left(n^{2-\varepsilon}\right)$ distinct triangles. (hint: There is no need to change the rich points proof — only the ESGK framework from Section 9.1). Also prove that this bound is tight up to the extra ε in the exponent.

Problem 9.4. Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^2$ that spans $O(n/\sqrt{\log n})$ distinct distances. Prove that there exists a rotation of $\mathbb{R}^2$ that takes $\Omega(2^{\log^{0.49} n})$ points of $\mathcal{P}$ to points of $\mathcal{P}$. Hint: Take another look at the calculation below the statement of Theorem 9.1.

9.5 Open problems

In this chapter we studied the number of rich points in a set of n lines in $\mathbb{R}^3$, when every constant-degree two-dimensional variety contains $O(n^{1/2})$ of the lines. Claim 9.5 shows that our bound for this problem cannot be significantly improved. However, it is not clear what happens when we further restrict the lines.

Open Problem 9.1. Let $\mathcal{L}$ be a set of n lines in $\mathbb{R}^3$ such that every constant-degree two-dimensional variety in $\mathbb{R}^3$ contains $O(1)$ lines of $\mathcal{L}$. For $r \geq 2$, what is the maximum number of r-rich points that $\mathcal{L}$ can have?

Beyond being difficult to solve, this problem is also interesting since it is related to several other main open problems. For example, the unit distances problem (presented in Section 1.5) can be reduced to a variant of Open Problem 9.1. One may also ask what happens when at most q lines are contained in a constant-degree two-dimensional variety, for any $q = o(n^{1/2})$.

As already stated in Chapter 1, the distinct distances problem was asked for point sets in $\mathbb{R}^d$. That is, what is the minimum number of distinct distances that can be determined by a set of n points in $\mathbb{R}^d$. When $d \geq 3$, the gap between the current best lower and upper bounds is polynomial in n. For such d, it is not difficult to show that an $n^{1/d} \times n^{1/d} \times \cdots \times n^{1/d}$ section of the integer lattice $\mathbb{Z}^d$ spans $\Theta(n^{2/d})$ distances, and this is conjectured to be tight.

Conjecture 9.11. *For $d \geq 3$, every set of n points in $\mathbb{R}^d$ determines $\Omega(n^{2/d})$ distinct distances.*

The current best bound is a combination of a result of Solymosi and Vu [91] with the planar bound of Guth and Katz [51]. For the full details, see for example [85].

One common belief is that the approach of Guth and Katz in $\mathbb{R}^2$ could be extended to $\mathbb{R}^d$. Bardwell-Evans and Sheffer [4] reduced the distinct distances problem in $\mathbb{R}^d$ to an incidence problem with $(d-1)$-flats in $\mathbb{R}^{2d-1}$. These flats are well-behaved in several ways, such as that every two flats intersect in at most one point. However, it is currently not known how to solve such incidence problems.

Chapter 10

Distinct Distances Variants

After a rather technical chapter which completed the proof of the distinct distances theorem, we move to a short lighter chapter. We now study the current best bounds for a couple of open variants of the distinct distances problem.

10.1 Subsets with no repeated distances

Given a set $\mathcal{P}$ of points in $\mathbb{R}^2$, let $\mathsf{subset}(\mathcal{P})$ denote the size of the largest subset $\mathcal{P}' \subset \mathcal{P}$ such that every distance is spanned by the points of $\mathcal{P}'$ at most once. That is, there are no points $a, b, c, d \in \mathcal{P}'$ such that $|ab| = |cd| > 0$ (including cases where $a = c$). Figure 10.1 depicts a set of 25 points and a subset of four points that span every distance at most once.
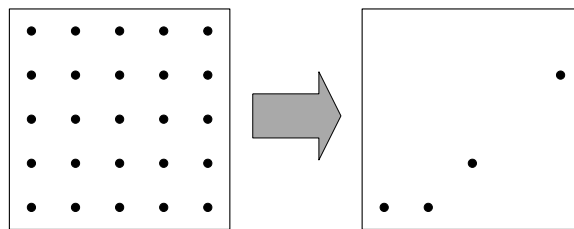
Figure 10.1: A set of 25 points and a subset of four points that span every distance at most once. No subset of five points has this property.

Let $\mathsf{subset}(n) = \min_{|\mathcal{P}|=n} \mathsf{subset}(\mathcal{P})$. In other words, $\mathsf{subset}(n)$ is the maximum number satisfying that every set of n points in $\mathbb{R}^2$ contains a subset of $\mathsf{subset}(n)$ points that do not span any distance more than once. We are interested in the asymptotic value of $\mathsf{subset}(n)$.

117

If every distance spanned by a subset $\mathcal{P}' \subset \mathcal{P}$ is unique, then the number of distances spanned by $\mathcal{P}$ must be at least $\binom{|P'|}{2}$. That is, when $\mathcal{P}$ spans d distinct distances we have $|\mathcal{P}'| = O\left(\sqrt{d}\right)$. Let $\mathcal{L}$ be a $\sqrt{n} \times \sqrt{n}$ section of the integer lattice $\mathbb{Z}^2$. By Theorem 1.9, the set $\mathcal{L}$ spans $\Theta(n/\sqrt{\log n})$ distinct distances. Therefore, we have $\mathsf{subset}(n) \leq \mathsf{subset}(\mathcal{L}) = O\left(\sqrt{n}/\log^{1/4} n\right)$. This is the current best upper bound for $\mathsf{subset}(n)$.

The following is the current best lower bound for $\mathsf{subset}(n)$.

Theorem 10.1 (Charalambides [18]). $\mathsf{subset}(n) = \Omega\left(n^{1/3}/\log^{1/3} n\right)$.

The proof of Theorem 10.1 combines the Guth-Katz distinct distances result with a simple probabilistic argument. For this proof, we require a bound on the maximum number of isosceles triangles that are spanned by a set of n points in $\mathbb{R}^2$ (triangles whose three vertices are in the point set). The current best bound, by Pach and Tardos [71], is $O(n^{2.137})$. The following weaker bound suffices for proving Theorem 10.1.

Claim 10.2. *Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^2$. Then $\mathcal{P}$ determines $O(n^{7/3})$ isosceles triangles.*

Proof. For a fixed point $p \in \mathcal{P}$, we bound the number of isosceles triangles having p as one of the vertices incident to the base edge. For a point $q \in \mathcal{P} \setminus \{p\}$, we consider such triangles where q is the vertex not incident to the base edge. The third vertex of such a triangle is on the circle centered at q and incident to p (see Figure 10.2(a)). We denote this circle as C_q, and set $\mathcal{C} = \{C_q : q \in \mathcal{P} \setminus \{p\}\}$. The number of isosceles triangles in which p is incident to the base edge is $I(\mathcal{P} \setminus \{p\}, \mathcal{C})$.

Note that no two circles of $\mathcal{C}$ are identical, since each circle has a distinct center. In a general point-circle incidence problem, the incidence graph may contain a $K_{2,2}$. In the current scenario, since the circles of $\mathcal{C}$ are all incident to p, the incidence graph of $(\mathcal{P} \setminus \{p\}) \times \mathcal{C}$ does not contain a $K_{2,2}$ (see Figure 10.2(b)). Theorem 3.3 implies that $I(\mathcal{P} \setminus \{p\}, \mathcal{C}) = O\left(n^{4/3}\right)$. By summing this bound over every $p \in \mathcal{P}$, we obtain that the number of isosceles triangles spanned by $\mathcal{P}$ is $O(n^{7/3})$. $\square$

We are now ready to derive an upper bound for $\mathsf{subset}(n)$.

Proof of Theorem 10.1. Consider a set $\mathcal{P}$ of n points in $\mathbb{R}^2$. Similarly to the ESGK framework, we define the set

$$Q = \left\{(a, b, c, d) \in \mathcal{P}^4 \; : \; |ab| = |cd| > 0\right\}.$$
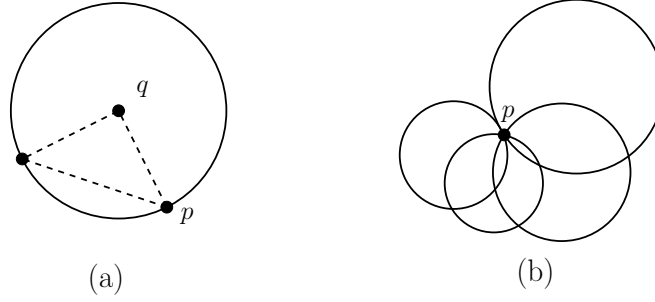
Figure 10.2: (a) In an isosceles triangle where p incident to the base edge but not q, the third vertex is on the circle centered at q and incident to p. (b) Circles that intersect at p lead to an incidence graph with no $K_{2,2}$ (when p is not a vertex of this graph).

Unlike the ESGK framework, we only consider quadruples that consist of four distinct points. In Chapters 8 and 9, we proved that $|Q| = O(n^{3+\varepsilon})$. The more involved proof of Guth and Katz implies $|Q| = O(n^3 \log n)$. Let T be the set of isosceles triangles spanned by points of $\mathcal{P}$ (including equilateral triangles). By Claim 10.2, we have $|T| = O(n^{7/3})$.

Consider a probability $0 < p < 1$ that will be set below. Let $\mathcal{P}' \subseteq \mathcal{P}$ be a subset that is obtained by choosing every point of $\mathcal{P}$ with probability p. Then we have the expectation $\mathbb{E}[|\mathcal{P}'|] = pn$. Let $Q' \subseteq Q$ be the set of quadruples of Q that contain only points of $\mathcal{P}'$. Every quadruple of Q is in Q' with a probability of p^4, so $\mathbb{E}[|Q'|] \le \alpha p^4 n^3 \log n$ for a sufficiently large constant α. Let $T' \subseteq T$ be the set of isosceles triangles of T that have their three vertices in $\mathcal{P}'$. We have $\mathbb{E}[|T'|] \le \alpha p^3 n^{7/3}$, for a sufficiently large constant α. The points of $\mathcal{P}'$ span every distance at most once if and only if $|Q'| = |T'| = 0$. We fix a sufficiently large constant α that satisfies the above. By linearity of expectation we have

$$\mathbb{E}\left[|\mathcal{P}'| - |Q'| - |T'|\right] \ge pn - \alpha p^4 n^3 \log n - \alpha p^3 n^{7/3}.$$

By setting $p = 1/(2\alpha n^2 \log n)^{1/3}$, for sufficiently large n we obtain

$$\mathbb{E}\left[|\mathcal{P}'| - |Q'| - |T'|\right] \ge \frac{n^{1/3}}{2^{1/3}\alpha^{1/3}\log^{1/3} n} - \frac{n^{1/3}}{2^{4/3}\alpha^{1/3}\log^{1/3} n} - \frac{n^{1/3}}{2\log n} > \frac{n^{1/3}}{3\alpha^{1/3}\log^{1/3} n}.$$

Thus, there exists a subset $\mathcal{P}' \subset \mathcal{P}$ for which $|\mathcal{P}'| - |Q'| - |T'| > \frac{n^{1/3}}{3(\alpha_1 \log n)^{1/3}}$. We construct $\mathcal{P}'' \subseteq \mathcal{P}'$ by arbitrarily removing from $\mathcal{P}'$ a point from every tuple of Q' and T'. The subset $\mathcal{P}''$ does not span any repeated distances and contains $\Omega(n^{1/3}/\log^{1/3} n)$ points of $\mathcal{P}$. $\qquad\square$

10.2   Point sets with few distinct distances

Characterizing the point sets in $\mathbb{R}^2$ that determine a small number of distinct distances appears to be one of the most difficult open problems concerning distinct distances. Erdős asked whether every set of n points in $\mathbb{R}^2$ that spans $O(n/\sqrt{\log n})$ distances "has lattice structure" [37]. In Chapter 1 we saw that a $\sqrt{n} \times \sqrt{n}$ section of $\mathbb{Z}^2$ determines $O(n/\sqrt{\log n})$ distinct distances. The same holds for $\sqrt{n} \times \sqrt{n}$ sections of other lattices (for example, see [85]). Since proving the conjecture seems to be very difficult, Erdős suggested to first prove that for every such point set there exists a line containing $\Omega(\sqrt{n})$ points of the set [37]. This would imply that the point set can be covered by a relatively small number of lines. Since this also appears to be quite difficult, Erdős asked whether there exists a line containing $\Omega(n^\varepsilon)$ points of the set, for any $\varepsilon > 0$. Embarrassingly, even this weaker variant remains open after several decades.

We now present the current best bound for the above problem. To derive this bound, we require the following straightforward generalization of Theorem 3.3 (see also Problem 3.14).

Theorem 10.3. *Let $\mathcal{P}$ be a set of m points and let Γ be a set of n distinct irreducible algebraic curves of degree at most k, both in $\mathbb{R}^2$. If the incidence graph of $\mathcal{P} \times \Gamma$ contains no copy of $K_{s,t}$, then*

$$I(\mathcal{P}, \Gamma) = O_{s,k}\left(m^{\frac{s}{2s-1}} n^{\frac{2s-2}{2s-1}} t^{\frac{1}{2s-1}} + tm + n \right).$$

Claim 10.4. *Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^2$ that spans $O(n/\sqrt{\log n})$ distinct distances. Then there exists a line that contains $\Omega(\log n)$ points of $\mathcal{P}$.*

Proof. Let D be the set of distances spanned by $\mathcal{P}$. We construct a set $\mathcal{C}$ of circles by placing $|D| = O(n/\sqrt{\log n})$ circles around every point of $\mathcal{P}$. For each point, we place one circle for each distance of D, where the distance is the radius of the corresponding circle. Note that $|\mathcal{C}| = O(n^2/\sqrt{\log n})$. Since the circles centered at a point $p \in \mathcal{P}$ form exactly one incidence with each point of $\mathcal{P} \setminus \{p\}$, we have

$$I(\mathcal{P}, \mathcal{C}) = n(n-1). \tag{10.1}$$

Let x denote the maximum number of points of $\mathcal{P}$ that are on a common line. Given two points $p, q \in \mathcal{P}$, if a circle $C \in \mathcal{C}$ contains both p and q then the perpendicular bisector of p and q is incident to the center of C. This implies that the incidence

graph of $\mathcal{P} \times \mathcal{C}$ contains no copy of $K_{2,x+1}$. Applying Theorem 10.3 implies

$$I(\mathcal{P}, \mathcal{C}) = O\left(n^{2/3}\left(\frac{n^2}{\sqrt{\log n}}\right)^{2/3} x^{1/3} + nx + \frac{n^2}{\sqrt{\log n}}\right)$$

$$= O\left(\frac{n^2 x^{1/3}}{\log^{1/3} n} + \frac{n^2}{\sqrt{\log n}} + nx\right).$$

Combining this with (10.1) implies the assertion of the claim. $\qquad\square$

10.3 Exercises

Problem 10.1. Let $A \subset \mathbb{R}$ be a set of n real numbers. Prove that there exists a subset $A' \subset A$ such that $|A'| = \Omega(n^{1/3})$ and no *difference* repeats more than once in A'. That is, there do not exist $a, b, c, d \in A'$ such that $a - b = c - d$. (Hint: Define a set of quadruples and derive an upper bound on its size.)

Problem 10.2. Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^2$, such that no four distinct points $a, b, c, d \in \mathcal{P}$ are the vertices of an isosceles trapezoid. Prove that $\mathcal{P}$ determines $\Omega(n)$ distinct distances. You might like to use the following approach.

Let x be the number of distinct distances spanned by $\mathcal{P}$. Consider the set of isosceles triangles that are spanned by $\mathcal{P}$:

$$T = \{(a, b, c) \in \mathcal{P}^3 : |ab| = |ac| \text{ and } b \neq c\}. \tag{10.2}$$

Prove the claim by double counting $|T|$. Use the Cauchy-Schwarz inequality to find a lower bound for the number of triangles that involve a specific $a \in \mathcal{P}$. The points of $\mathcal{P} \setminus \{a\}$ are on at most x circles centered at a. Two points $b, c \in \mathcal{P}$ form a triple with a if and only if they are on the same circle. To derive an upper bound for $|T|$, find a connection to perpendicular bisectors of pairs of points of $\mathcal{P}$.

Problem 10.3. Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^2$, such that no three points of $\mathcal{P}$ are collinear. Prove that there exists a point $p \in \mathcal{P}$ such that the number of distinct distances between p and $\mathcal{P} \setminus \{p\}$ is at least $(n-1)/3$.

You might like to imitate the proof of Problem 10.2, by defining T as in (10.2) and double counting $|T|$. In this case, x is the maximum number of distinct distances between any point $p \in \mathcal{P}$ and $\mathcal{P} \setminus \{p\}$.

Chapter 11

Incidences in $\mathbb{R}^d$

In this chapter we continue to study incidences in $\mathbb{R}^d$. As a warmup, we derive an incidence bound for curves in $\mathbb{R}^3$ (stronger than the bound presented in Chapter 7). The main result of this section is a general point-variety incidence bound in $\mathbb{R}^d$. Deriving this result requires additional tools from Algebraic Geometry, and in particular the concept of Hilbert polynomials.

11.1 Warmup: Incidences with curves in $\mathbb{R}^3$

In Chapter 7 we saw how to use the "constant-degree polynomial partitioning" technique to derive incidence bounds in $\mathbb{R}^d$. In particular, Theorem 7.5 contains a point-variety incidence bound that holds in any dimension d. Recall that this theorem applies only to incidences with varieties of dimension at most $d/2$. Moreover, it seems to give the conjectured incidence bound only when the dimension of the varieties is exactly $d/2$ (see Section 11.5 for details about the conjectured bound). As a first example of improving the bound of Theorem 7.5, we now derive a stronger bound for curves in $\mathbb{R}^3$. The comparison to Theorem 7.5 is not accurate, since the assumptions satisfied by the curves are rather different in the two cases. The following bound has a simple proof, and is presented as a warmup before getting to a more involved technique for incidences in $\mathbb{R}^d$. A reader who does not wish to see yet another use of constant-degree polynomial partitioning may prefer to skip this section.

We recall Lemma 7.2, in which we proved that our point-curve incidence bound in $\mathbb{R}^2$ holds in any dimension.

Lemma 7.2. *Consider an integer $d \geq 2$. Let $\mathcal{P}$ be a set of m points and let Γ be a set of n varieties of dimension at most one and degree at most k, both in $\mathbb{R}^d$. If the*

incidence graph of $\mathcal{P} \times \Gamma$ *contains no copy of* $K_{s,t}$, *then*

$$I(\mathcal{P}, \Gamma) = O_{s,t,k,d}\left(m^{\frac{s}{2s-1}}n^{\frac{2s-2}{2s-1}} + m + n\right).$$

Without additional assumptions we cannot expect a point-curve bound in $\mathbb{R}^3$ (or in any dimension $d \geq 3$) to be stronger than the point-curve incidence bound in $\mathbb{R}^2$. Indeed, assume that a point set $\mathcal{P}$ and a set of curves Γ in $\mathbb{R}^2$ yield a large number of incidences. We can obtain the same number of incidences in $\mathbb{R}^3$ by taking an arbitrary plane $\Pi \subset \mathbb{R}^3$ and placing $\mathcal{P}$ and Γ in Π. For the same reason, any general point-curve incidence bound that holds in $\mathbb{R}^d$ for some $d \geq 3$ also holds in $\mathbb{R}^2$.

In the above explanation for why we cannot obtain a stronger point-curve incidence bound in $\mathbb{R}^3$, our example is two-dimensional: The curves are contained in a single plane. When dealing with specific types of curves we can sometimes replace the plane with a different surface. For example, it is possible to take a configuration of circles in $\mathbb{R}^2$ and place it on a sphere in $\mathbb{R}^3$ while maintaining incidences. If the set of curves is required to be "truly three-dimensional" in the sense that no low-degree surface contains many curves, then we can obtain a stronger bound.

Theorem 11.1. *For any* $\varepsilon > 0$ *there exists a constant* c_ε *that satisfies the following. Let* $\mathcal{P}$ *be a set of* m *points and let* Γ *be a set of* n *irreducible algebraic curves of degree at most* k, *both in* $\mathbb{R}^3$. *Assume that the incidence graph of* $\mathcal{P} \times \Gamma$ *contains no copy of* $K_{s,t}$ *and that every two-dimensional variety in* $\mathbb{R}^3$ *of degree at most* c_ε *contains at most* q *curves of* Γ. *Then*

$$I(\mathcal{P}, \Gamma) = O_{s,t,k,\varepsilon}\left(m^{\frac{s}{3s-2}+\varepsilon}n^{\frac{3s-3}{3s-2}} + m^{\frac{s}{2s-1}+\varepsilon}n^{\frac{3s-3}{4s-2}}q^{\frac{s-1}{4s-2}} + m + n\right).$$

As an example, consider the case where Γ is a set of lines, $m = n$, and every surface of degree at most c_ε contains at most $\sqrt{n}$ lines of Γ. Applying Theorem 11.1 with $s = 2$ and $q = \sqrt{n}$ gives $I(\mathcal{P}, \Gamma) = O_\varepsilon\left(n^{5/4+\varepsilon}\right)$. Lemma 7.2 leads to the weaker bound $I(\mathcal{P}, \Gamma) = O\left(n^{4/3}\right)$.

Proof of Theorem 11.1. Let α_1 and α_2 be sufficiently large constants that depend on s, t, k, and ε. The hidden constants in the $O(\cdot)$-notations throughout the proof may also depend on s, t, k, and ε. For brevity we write $O(\cdot)$ instead of $O_{s,t,k,\varepsilon}(\cdot)$. We prove by induction on $m + n$ that

$$I(\mathcal{P}, \Gamma) \leq \alpha_1\left(m^{\frac{s}{3s-2}+\varepsilon}n^{\frac{3s-3}{3s-2}} + m^{\frac{s}{2s-1}+\varepsilon}n^{\frac{3s-3}{4s-2}}q^{\frac{s-1}{4s-2}}\right) + \alpha_2(m+n).$$

For the induction base, the case where m and n are sufficiently small can be handled by choosing sufficiently large values of α_1 and α_2. We move to consider

the induction step. Since the incidence graph contains no copy of $K_{s,t}$, Theorem 7.1 implies $I(\mathcal{P}, \Gamma) = O(mn^{1-1/s} + n)$. When $m = O(n^{1/s})$ this implies $I(\mathcal{P}, \Gamma) = O(n)$, and we may thus assume that

$$n = O(m^s). \tag{11.1}$$

Partitioning the space.　Let f be an r-partitioning polynomial of $\mathcal{P}$, for a sufficiently large constant r. The asymptotic relations between the various constants are

$$2^{1/\varepsilon}, k, s, t \ll r \ll \alpha_2 \ll \alpha_1.$$

By the polynomial partitioning theorem (Theorem 3.1), we have $\deg f = O(r)$. By Warren's theorem (Theorem 3.2), the number of cells is $c = O(r^3)$. Denote the cells of the partition as $C_1, \ldots, C_c$. For each $j = 1, \ldots, c$, let Γ_j denote the set of curves of Γ that intersect C_j and set $\mathcal{P}_j = C_j \cap \mathcal{P}$. We also set $m_j = |\mathcal{P}_j|$, $m' = \sum_{j=1}^c m_j$, and $n_j = |\Gamma_j|$. Note that $m_j \leq m/r^3$ for every $1 \leq j \leq c$. For any curve $\gamma \in \Gamma$, Theorem 4.11 with $U = \gamma$ and $W = \mathbf{V}(f)$ implies that γ intersects $O(r)$ cells of $\mathbb{R}^3 \setminus \mathbf{V}(f)$. Therefore, $\sum_{j=1}^c n_j = O(nr)$, and according to Hölder's inequality we have

$$\sum_{j=1}^c n_j^{\frac{3s-3}{3s-2}} \leq \left(\sum_{j=1}^c n_j \right)^{\frac{3s-3}{3s-2}} \left(\sum_{j=1}^c 1 \right)^{\frac{1}{3s-2}} = O\left((nr)^{\frac{3s-3}{3s-2}} r^{\frac{3}{3s-2}} \right) = O\left(n^{\frac{3s-3}{3s-2}} r^{\frac{3s}{3s-2}} \right),$$

$$\sum_{j=1}^c n_j^{\frac{3s-3}{4s-2}} \leq \left(\sum_{j=1}^c n_j \right)^{\frac{3s-3}{4s-2}} \left(\sum_{j=1}^c 1 \right)^{\frac{s+1}{4s-2}} = O\left((nr)^{\frac{3s-3}{4s-2}} r^{\frac{3s+3}{4s-2}} \right) = O\left(n^{\frac{3s-3}{4s-2}} r^{\frac{3s}{2s-1}} \right).$$

Combining the above with the induction hypothesis implies

$$\sum_{j=1}^c I(\mathcal{P}_j, \Gamma_j) \leq \sum_{j=1}^c \left(\alpha_1 \left(m_j^{\frac{s}{3s-2}+\varepsilon} n_j^{\frac{3s-3}{3s-2}} + m_j^{\frac{s}{2s-1}+\varepsilon} n_j^{\frac{3s-3}{4s-2}} q^{\frac{s-1}{4s-2}} \right) + \alpha_2 (m_j + n_j) \right)$$

$$\leq \alpha_1 \left(\frac{m^{\frac{s}{3s-2}+\varepsilon}}{r^{\frac{3s}{3s-2}+3\varepsilon}} \sum_{j=1}^c n_j^{\frac{3s-3}{3s-2}} + \frac{m^{\frac{s}{2s-1}+\varepsilon} q^{\frac{s-1}{4s-2}}}{r^{\frac{3s}{2s-1}+3\varepsilon}} \sum_{j=1}^c n_j^{\frac{3s-3}{4s-2}} \right) + \sum_{j=1}^c \alpha_2 (m_j + n_j)$$

$$= \alpha_1 \cdot O \left(\frac{m^{\frac{s}{3s-2}+\varepsilon} n^{\frac{3s-3}{3s-2}}}{r^{3\varepsilon}} + \frac{m^{\frac{s}{2s-1}+\varepsilon} n^{\frac{3s-3}{4s-2}} q^{\frac{s-1}{4s-2}}}{r^{3\varepsilon}} \right) + \alpha_2 \left(m' + O(nr) \right).$$

By (11.1) we have $n = O\left(m^{\frac{s}{3s-2}} n^{\frac{3s-3}{3s-2}} \right)$. Thus, when α_1 is sufficiently large with

respect to r and α_2, we get

$$\sum_{j=1}^{c} I(\mathcal{P}_j, \Gamma_j) = \alpha_1 \cdot O\left(\frac{m^{\frac{s}{3s-2}+\varepsilon}n^{\frac{3s-3}{3s-2}}}{r^{3\varepsilon}} + \frac{m^{\frac{s}{2s-1}+\varepsilon}n^{\frac{3s-3}{4s-2}}q^{\frac{s-1}{4s-2}}}{r^{3\varepsilon}}\right) + \alpha_2 m'.$$

When r is sufficiently large with respect to ε and to the constant hidden in the $O(\cdot)$-notation, we have

$$\sum_{j=1}^{c} I(\mathcal{P}_j, \Gamma_j) \le \frac{\alpha_1}{2}\left(m^{\frac{s}{3s-2}+\varepsilon}n^{\frac{3s-3}{3s-2}} + m^{\frac{s}{2s-1}+\varepsilon}n^{\frac{3s-3}{4s-2}}q^{\frac{s-1}{4s-2}}\right) + \alpha_2 m'. \tag{11.2}$$

Incidences on the partition. It remains to study incidences with points that lie on $\mathbf{V}(f)$. Set $\mathcal{P}_0 = \mathcal{P} \cap \mathbf{V}(f)$ and $m_0 = |\mathcal{P}_0| = m - m'$. Let Γ_0 denote the set of curves that are contained in $\mathbf{V}(f)$. Set $\Gamma' = \Gamma \setminus \Gamma_0$, $n_0 = |\Gamma_0|$, and $n' = |\Gamma'| = n - n_0$. By Theorem 4.11 every curve of Γ' intersects $\mathbf{V}(f)$ in $O(r)$ points. That is, every such curve is incident to $O(r)$ points of $\mathcal{P}_0$. Summing this over every curve of Γ' and taking α_2 to be sufficiently large gives

$$I(\mathcal{P}_0, \Gamma') = O(nr) \le \frac{\alpha_2}{2}n. \tag{11.3}$$

It remains to derive a bound for $I(\mathcal{P}_0, \Gamma_0)$. For this purpose, we set c_ε so that $c_\varepsilon > \deg f$ and $c_\varepsilon = O(r)$. By the assumption of the theorem we get that $|\Gamma_0| \le q$. By applying Lemma 7.2 (which is restated in the beginning of this section) on $\mathcal{P}_0$ and Γ_0, we obtain

$$I(\mathcal{P}_0, \Gamma_0) = O\left(m_0^{\frac{s}{2s-1}}q^{\frac{2s-2}{2s-1}} + m_0 + q\right).$$

Since $q \le n$ and $m_0 \le m$, we have

$$m_0^{\frac{s}{2s-1}}q^{\frac{2s-2}{2s-1}} \le m^{\frac{s}{2s-1}}n^{\frac{3s-3}{4s-2}}q^{\frac{s-1}{4s-2}}.$$

Combining this with a sufficiently large choice of α_1 and α_2 implies

$$I(\mathcal{P}_0, \Gamma_0) = O\left(m^{\frac{s}{2s-1}}n^{\frac{3s-3}{4s-2}}q^{\frac{s-1}{4s-2}} + n + m_0\right)$$

$$\le \frac{\alpha_1}{2}m^{\frac{s}{2s-1}}n^{\frac{3s-3}{4s-2}}q^{\frac{s-1}{4s-2}} + \frac{\alpha_2}{2}(n + m_0). \tag{11.4}$$

By combining (11.2), (11.3), and (11.4), we complete the induction step and thus the proof of the theorem. □

It is not difficult to extend Theorem 11.1 to incidences with curves in any dimension. As the dimension grows, the bound becomes stronger while the restrictions over the set of curves become more involved (see [86]). That is, we can improve Theorem 7.5 when $d \geq 3$ and the varieties are curves. This improved bound is obtained under different restrictions on the curves: Instead of a restriction concerning the tangents of the curves we have a restriction concerning the number of curves in a low-degree variety. It is conjectured that similar improved bounds exist whenever the dimension of the varieties is not $d/2$ (see Section 11.5 for the conjectured bounds).

So far the conjectured bounds for incidences in $\mathbb{R}^d$ have only been obtained for curves, varieties of dimension $d/2$, and hypersurfaces. We have seen the case of curves in Theorem 11.1 and the case of varieties of degree $d/2$ in Theorem 7.5. In the following sections we will study the case of hypersurfaces.

11.2   Hilbert polynomials

As already stated, the main goal of this chapter is to derive a general point-variety incidence bound in $\mathbb{R}^d$. This bound is derived in Section 11.3. In the current section we prepare for the proof in Section 11.3 by introducing more tools from Algebraic Geometry.

For an integer $m \geq 0$, we denote by $\mathbb{R}[x_1, \ldots, x_d]_{\leq m}$ the set of polynomials of degree at most m in $\mathbb{R}[x_1, \ldots, x_d]$. Similarly, if $J \subset \mathbb{R}[x_1, \ldots, x_d]$ is an ideal, we denote by $J_{\leq m} = J \cap \mathbb{R}[x_1, \ldots, x_d]_{\leq m}$ the set of polynomials in J of degree at most m. Note that $\mathbb{R}[x_1, \ldots, x_d]_{\leq m}$ is not a ring and that $J_{\leq m}$ is not an ideal. As seen in the proof of Theorem 3.6, there are $\binom{d+m}{m}$ monomials in the variables $x_1, \ldots, x_d$ of degree at most m (ignoring the coefficient of the monomial). Thus, we can consider $\mathbb{R}[x_1, \ldots, x_d]_{\leq m}$ as a vector space of dimension $\binom{d+m}{m}$. Specifically, $\mathbb{R}[x_1, \ldots, x_d]_{\leq m}$ is isomorphic to $\mathbb{R}^{\binom{d+m}{m}}$.[1] Similarly, $J_{\leq m}$ is a vector space of a finite dimension (for example, see Problem 11.1(a)) and a subspace of $\mathbb{R}[x_1, \ldots, x_d]_{\leq m}$.

As a simple example, consider the ideal $J = \langle x^2y, xy^2 \rangle \subset \mathbb{R}[x, y]$ (that is, every term of every polynomial in J is a multiple of x^2y or of xy^2). An ideal that is generated by monomials is called *a monomial ideal*. While all of the following definitions and results hold for arbitrary ideals, we use a monomial ideal as an example since it is

---

[1]Since constant multiples of the same polynomial define the same variety, we can consider such multiples as equivalent. With this equivalence relation $\mathbb{R}[x_1, \ldots, x_d]_{\leq m}$ behaves like the projective space $\mathbb{R}\mathbf{P}^{\binom{d+m}{m}-1}$. We will not rely on any projective properties in this chapter, so we stick to the affine space $\mathbb{R}^{\binom{d+m}{m}}$.
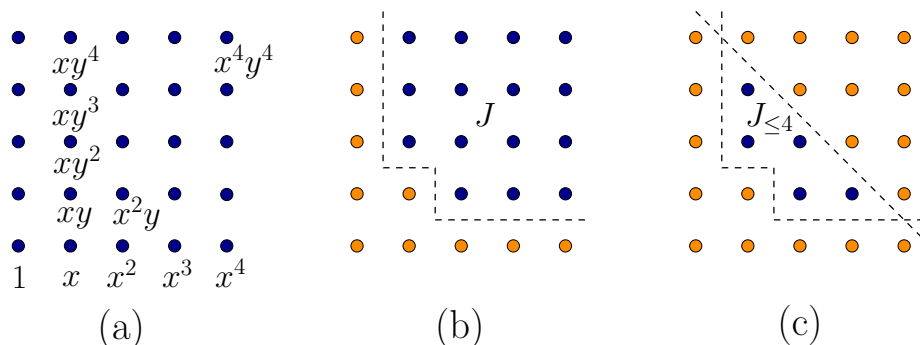
Figure 11.1: (a) Every lattice point corresponds to a monomial of $\mathbb{R}[x, y]$. (b) The ideal $J = \langle x^2 y, xy^2 \rangle$. (c) The vector space $J_{\leq 4}$.

easier to understand. The first advantage of dealing with a monomial ideal as an example is that we can think of the set of monomials in $\mathbb{R}[x, y]$ as a lattice (ignoring the real coefficients), as shown in Figure 11.1(a). Figure 11.1(b) shows the structure of J, and Figure 11.1(c) shows $J_{\leq 4}$.

Every element $\mathbb{R}[x, y]_{\leq 3}$ can be written as

$$c_1 + c_2 x + c_3 y + c_4 x^2 + c_5 xy + c_6 y^2 + c_7 x^3 + c_8 x^2 y + c_9 xy^2 + c_{10} y^3,$$

and we can think of it as the point $(c_1, \ldots, c_{10}) \in \mathbb{R}^{10}$. We write the coordinates of $\mathbb{R}^{10}$ as $x_1, \ldots, x_{10}$. In the example ideal $J = \langle x^2 y, xy^2 \rangle \subset \mathbb{R}[x, y]$, we get that $J_{\leq 3}$ is the 2-flat in $\mathbb{R}^{10}$ defined by $x_1 = \cdots = x_7 = x_{10} = 0$.

The quotient $\mathbb{R}[x_1, \ldots, x_d]_{\leq m}/J_{\leq m}$ is also a vector space.[2] The *Hilbert function* of an ideal $J \subset \mathbb{R}[x_1, \ldots, x_d]$ is defined as

$$h_J(m) = \dim \left(\mathbb{R}[x_1, \ldots, x_d]_{\leq m}/J_{\leq m} \right).$$

In the example $J = \langle x^2 y, xy^2 \rangle \subset \mathbb{R}[x, y]$, the quotient $\mathbb{R}[x, y]_{\leq m}/J_{\leq m}$ is the set of polynomials in $\mathbb{R}[x, y]_{\leq m}$ that do not have monomials that are divisible by $x^2 y$ and xy^2. Equivalently, $\mathbb{R}[x, y]_{\leq m}/J_{\leq m}$ is the set of polynomials in $\mathbb{R}[x, y]_{\leq m}$ that are generated by monomials that are not in J (this property is special to monomial ideals, and does not hold in general). When $m = 3$, the quotient $\mathbb{R}[x, y]_{\leq 3}/J_{\leq 3}$ is the set of polynomials of the form $c_1 + c_2 x + c_3 y + c_4 x^2 + c_5 xy + c_6 y^2 + c_7 x^3 + c_{10} y^3$. More

---

[2]Given a vector space A and a subspace B, recall that in the *quotient* A/B two elements $a, a' \in A$ are equivalent if and only if there exists $b \in B$ such that $a + b = a'$.

generally, we get that

$$h_J(m) = \dim\left(\mathbb{R}[x,y]_{\leq m}/J_{\leq m}\right) = \begin{cases} 1, & m = 0, \\ 3, & m = 1, \\ 2 + 2m, & m \geq 2. \end{cases}$$

In the above, when ignoring the first two values of $h_J(m)$ we get that it is a linear polynomial. This is an example of the following property of Hilbert functions. For every ideal $J \subset \mathbb{R}[x_1, \ldots, x_d]$, there exists an integer $m_J \geq 0$ and a polynomial $H_J \in \mathbb{R}[m]$ such that for every $m > m_J$ we have $h_J(m) = H_J(m)$. That is, for sufficiently large m the Hilbert function behaves like a polynomial. The polynomial H_J is called the *Hilbert polynomial of J*, and m_J is called the *regularity* of J. From the above, we have that the Hilbert polynomial of $\langle x^2 y, xy^2 \rangle$ is $2 + 2m$ and that its regularity is 1.

Let $U \subset \mathbb{R}^d$ be a variety. Then the dimension of U is equal to the degree of the Hilbert polynomial $H_{\mathbf{I}(U)}(m)$. That is, Hilbert polynomials provide an alternative way to define the dimension of a variety. The coefficient of the leading term $m^{\dim U}$ of $H_{\mathbf{I}(U)}(m)$ is always positive. As with previous uses of an ideal of a variety $\mathbf{I}(U)$, the above definition of dimension does not remain valid for every ideal $J \subset \mathbb{R}[x_1, \ldots, x_d]$ that satisfies $U = \mathbf{V}(J)$. To get the correct dimension of U, one has to use the Hilbert polynomial of the ideal $\mathbf{I}(U)$.

Let $U \subset \mathbb{R}^d$ be a variety of degree k and dimension d', and let $J = \mathbf{I}(U)$. Then the regularity m_J of J is $O_{d,k}(1)$. By combining this with the above properties of the Hilbert polynomial, we get that for every $m > m_J$

$$h_J(m) = \Theta_{d,k}\left(m^{d'}\right). \tag{11.5}$$

A more detailed introduction to Hilbert polynomials, including proofs for most of the above claims, can be found in [23, Chapter 9]. For the bound on the regularity m_J, see [44, Theorem B] and [42].

11.3   A general point-variety incidence bound

We are now ready to state our general incidence bound in $\mathbb{R}^d$. This result was derived in [42].

Theorem 11.2. *Let $\mathcal{P}$ be a set of m points and let $\mathcal{V}$ be a set of n varieties of degree at most k, both in $\mathbb{R}^d$. Assume that the incidences graph of $\mathcal{P} \times \mathcal{V}$ does not contain*

a copy of $K_{s,t}$. Then for every $\varepsilon > 0$, we have

$$I(\mathcal{P}, \mathcal{V}) = O_{k,s,t,d,\varepsilon}\left(m^{\frac{(d-1)s}{ds-1}+\varepsilon}n^{\frac{d(s-1)}{ds-1}} + m + n\right).$$

Note that Theorem 11.2 holds for varieties of any dimension. It leads to the current best bounds for varieties of dimension larger than $d/2$, although it achieves the conjectured bounds only for hypersurfaces. An improved bound for the case when the varieties of $\mathcal{V}$ are of dimension smaller than $d-1$ can be found in Problem 11.2.

We now present a non-rigorous explanation of our strategy for proving Theorem 11.2. Once again, we rely on the constant-degree polynomial partitioning technique. As always, the main difficulty with this technique is to bound the number of incidences on the partition (that is, incidences with points that are not in any cell). In Chapter 7 we handled these incidences by using an assumption about the tangent spaces of the varieties. In Theorem 11.1 we handled these incidences by using the assumption that not many curves can be contained in a low-degree surface. In Theorem 11.2 we have arbitrary varieties with no additional assumptions to rely on.

Let $U \subset \mathbb{R}^d$ be a variety that is our constant-degree partition. Let $\mathcal{P}_0 = \mathcal{P} \cap U$ be the set of points on this partition. To handle incidences with the points of $\mathcal{P}_0$ we will take a *second constant-degree partitioning polynomial* $f_2 \in \mathbb{R}[x_1, \ldots, x_d]$. The polynomial f_2 is not required to partition $\mathbb{R}^d$ but rather U. That is, the cell of the second partition are the connected components of $U \setminus \mathbf{V}(f_2)$ and we would like none of these cells to contain many points of $\mathcal{P}_0$.

Even if we can obtain a second partitioning polynomial and handle the incidences in the second set of cells, there might remain additional incidences to study. In particular, there may be many points on $U_2 = U \cap \mathbf{V}(f_2)$, and incidences with points on U_2 were not handled in any of the cells. To handle these incidences, we take a *third* partition which divides U_2 into cells. We continue performing partitioning steps, and after each such step the remaining incidences are contained in a lower-dimensional variety.

To use the above approach of multiple polynomial partitions, we need to derive a variant of the polynomial partitioning theorem (Theorem 3.1). We begin by deriving a variant of the ham sandwich theorem (Theorem 3.6), by relying on Hilbert polynomials. Recall from Chapter 3 that a polynomial $f : \mathbb{R}^d \to \mathbb{R}$ *bisects* a finite point set $\mathcal{P} \subset \mathbb{R}^d$ if each of the two sets $\{x \in \mathbb{R}^d : f(x) < 0\}$ and $\{x \in \mathbb{R}^d : f(x) > 0\}$ contains at most $|\mathcal{P}|/2$ points of $\mathcal{P}$.

Lemma 11.3. *Let $U \subset \mathbb{R}^d$ be an irreducible variety of dimension $d' \geq 1$ and degree k, and let $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_t$ be finite sets of points in $\mathbb{R}^d$. Then there exists $f \in \mathbb{R}[x_1, \ldots, x_d]$*

such that $f \notin \mathbf{I}(U)$, the polynomial f bisects each of the sets $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_t$, and $\deg f = O_{d,k}(t^{1/d'})$.

Proof. This proof is a variant of the proof of Theorem 3.6. Let $J = \mathbf{I}(U)$. As stated in Section 11.2, there exists a constant $m_J = O_{d,k}(1)$ such that (11.5) holds for every $m > m_J$. That is, for every such m the quotient $\mathbb{R}[x_1, \ldots, x_d]_{\leq m}/J_{\leq m}$ is a vector space of dimension $d_m = \Theta_{d,k}(m^{d'})$. Consider the minimum integer m that satisfies $d_m \geq t$. If $m > m_J$ then

$$t = \Theta_{d,k}\left(m^{d'}\right), \qquad \text{or equivalently} \qquad m = \Theta_{d,k}\left(t^{1/d'}\right).$$

If $m \leq m_J = O_{d,k}(1)$ then we also have that $m = O_{d,k}\left(t^{1/d'}\right)$, by taking the constant of the $O(\cdot)$-notation to be sufficiently large.

Let $p_1, \ldots, p_{d_m}$ be a basis for the vector space $\mathbb{R}[x_1, \ldots, x_d]_{\leq m}/J_{\leq m}$. Consider the map $\phi \colon \mathbb{R}^d \to \mathbb{R}^{d_m}$ defined by

$$\phi(x) = (p_1(x), \ldots, p_{d_m}(x)).$$

For every $1 \leq j \leq t$, let $\mathcal{P}'_j = \phi(\mathcal{P}_j) \subset \mathbb{R}^{d_m}$. For any points $u_1, u_2 \in U$, the quotient $\mathbb{R}[x_1, \ldots, x_d]_{\leq m}/J_{\leq m}$ contains a polynomial that vanishes on u_1 but not on u_2. For example, a generic hyperplane through u_1 is not incident to u_2 and not in $J_{\leq m}$. This implies that ϕ is injective on $U = \mathbf{V}(J)$, and thus that $|\mathcal{P}'_j| = |\mathcal{P}_j|$. Since $d_m \geq t$, the original ham sandwich theorem (Theorem 3.5) states that there exists a hyperplane $\Pi \subset \mathbb{R}^{d_m}$ that bisects each of the sets $\mathcal{P}'_1, \mathcal{P}'_2, \ldots, \mathcal{P}'_k$. We write the coordinates of $\mathbb{R}^{d_m}$ as $y_1, \ldots, y_{d_m}$, and note that $\Pi = \mathbf{V}(a_1 y_1 + \ldots + a_{d_m} y_{d_m})$ for some $a_1, \ldots, a_{d_m} \in \mathbb{R}$. That is, for each $1 \leq j \leq t$ we have

$$|\{y \in \mathcal{P}'_j \ : \ a_1 y_1 + \cdots + a_{d_m} y_{d_m} > 0\}| \leq |\mathcal{P}'_j|/2,$$

and

$$|\{y \in \mathcal{P}'_j \ : \ a_1 y_1 + \ldots + a_{d_m} y_{d_m} < 0\}| \leq |\mathcal{P}'_j|/2.$$

We set $f = a_1 p_1 + \cdots + a_{d_m} p_{d_m} \in \mathbb{R}[x_1, \ldots, x_d]$. By the definition of the basis elements p_j, the polynomial f is of degree at most m and is not in J. For a point $p \in \mathbb{R}^d$, note that $f(x) > 0$ if and only if $a_1 \phi(p)_1 + \ldots + a_{d_m} \phi(p)_{d_m} > 0$ (where $\phi(p)_j$ is the j'th coordinate of $\phi(p)$). Thus, for each $1 \leq j \leq d_m$ we have

$$|\{x \in \mathcal{P}_j \colon f(x) > 0\}| \leq |\mathcal{P}_j|/2 \quad \text{and} \quad |\{x \in \mathcal{P}_j \colon f(x) < 0\}| \leq |\mathcal{P}_j|/2.$$

This completes the proof of the lemma.   □

Note that the main change between the proof of Lemma 11.3 and the proof of the original polynomial ham sandwich (Theorem 3.6) is that the Veronese map was replaced with a somewhat different map. The Veronese map takes points of $\mathbb{R}^d$ to a space isomorphic to $\mathbb{R}[x_1, \ldots, x_d]_{\leq m}$, and we look for our bisecting polynomial in that space. The new map takes points of $\mathbb{R}^d$ to a space isomorphic to $\mathbb{R}[x_1, \ldots, x_d]_{\leq m}/J_{\leq m}$. The reason for moving to $\mathbb{R}[x_1, \ldots, x_d]_{\leq m}/J_{\leq m}$ is that the only polynomial in this quotient that vanishes on every point of U is 0. Indeed, every polynomial that vanishes on U is in $J_{\leq m}$ and thus equivalent to 0. Since we move to a subset of $\mathbb{R}[x_1, \ldots, x_d]_{\leq m}$, we need to take a larger value of m to get to a space of dimension at least t. We bounded this increase in m using a basic property of Hilbert polynomials

The following result will allow us to use multiple polynomial partitions.

Theorem 11.4. *Let $\mathcal{P}$ be a set of m points in $\mathbb{R}^d$ and let $U \subset \mathbb{R}^d$ be an irreducible variety of dimension d' and degree k. Then there exists an r-partitioning polynomial f of $\mathcal{P}$ such that $f \notin \mathbf{I}(U)$ and $\deg f = O_{d,k}(r^{d/d'})$.*

For our purpose, we only need to use Theorem 11.4 in the case where $\mathcal{P} \subset U$. We do not include this restriction in the statement of the theorem, since omitting it does not affect the proof.

Proof of Theorem 11.4. This proof is almost identical to the proof of Theorem 3.1. Let $c_{d,k}$ be the hidden constant in the $O(\cdot)$-notation of the bound on $\deg f$ in the statement of Lemma 11.3. That is, applying Lemma 11.3 on U implies the existence of a polynomial of degree at most $c_{d,k}t^{1/d'}$. Let $J = \mathbf{I}(U)$. We will show that there exists a sequence of polynomials $f_0, f_1, f_2, \ldots$ such that the degree of f_j is smaller than $c_{d,k}2^{(j+1)/d'}/(2^{1/d'} - 1)$, the polynomial f_j in not in J, and every connected component of $\mathbb{R}^d \setminus \mathbf{V}(f_j)$ contains at most $m/2^j$ points of $\mathcal{P}$. This would complete the proof since we can then choose $f = f_s$, where s is the minimum integer satisfying $2^s \geq r^d$.

We prove the existence of f_j by induction on j. For the induction basis we may take $f_0 = 1$, so we move to the induction step. By the induction hypothesis, there exists a polynomial f_j of degree smaller than $c_{d,k}2^{(j+1)/d'}/(2^{1/d'} - 1)$ such that $f_j \notin J$ and every connected component of $\mathbb{R}^d \setminus \mathbf{V}(f_j)$ contains at most $m/2^j$ points of $\mathcal{P}$. Since $|\mathcal{P}| = m$, the number t of connected components of of $\mathbb{R}^d \setminus \mathbf{V}(f_j)$ that contain more than $m/2^{j+1}$ points of $\mathcal{P}$ is smaller than 2^{j+1}. Let $\mathcal{P}_1, \ldots, \mathcal{P}_t \subset \mathcal{P}$ be the subsets of $\mathcal{P}$ that are contained in each of these connected components (that is, $|\mathcal{P}_1|, \ldots, |\mathcal{P}_t| > m/2^{j+1}$). By Lemma 11.3, there is a polynomial g_j of degree smaller than $c_d2^{(j+1)/d'}$ that simultaneously bisects $\mathcal{P}_1, \ldots, \mathcal{P}_t$. Let $f_{j+1} = f_j \cdot g_j$. Note that every connected component of $\mathbb{R}^d \setminus \mathbf{V}(f_j \cdot g_j)$ contains at most $m/2^{j+1}$ points of $\mathcal{P}$

and that $f_j \cdot g_j$ is a polynomial of degree smaller than

$$\frac{c_d 2^{(j+1)/d'}}{2^{1/d'}-1} + c_d 2^{(j+1)/d'} = c_d 2^{(j+1)/d} \cdot \left(\frac{1}{2^{1/d'}-1}+1\right) = \frac{c_d 2^{(j+2)/d'}}{2^{1/d'}-1}.$$

Since $f_j \notin J$ and $g_j \notin J$, we get that $f_{j+1} \notin J$. This completes the induction step and the proof of the theorem. $\qquad\square$

Now that we established the existence of multiple polynomial partitions, we are ready to use these to derive an incidence bound. Theorem 11.2 is an immediate corollary of the following result (by setting $U = \mathbb{R}^d$).

Theorem 11.5. *Let $U \subset \mathbb{R}^d$ be an irreducible variety of dimension d' and degree at most k. Let $\mathcal{P} \subset U$ be a set of m points and let $\mathcal{V}$ be a set of n varieties of degree at most k in $\mathbb{R}^d$, such that the incidences graph of $\mathcal{P} \times \mathcal{V}$ does not contain a copy of $K_{s,t}$. In addition, no variety of $\mathcal{V}$ contains U. Then for every $\varepsilon > 0$, we have*

$$I(\mathcal{P},\mathcal{V}) = O_{k,s,t,d,\varepsilon}\left(m^{\frac{(d'-1)s}{d's-1}+\varepsilon} n^{\frac{d'(s-1)}{d's-1}} + m + n\right).$$

Proof. We prove the theorem by induction on d'. Let $\alpha_{1,d',k}$ and $\alpha_{2,d',k}$ be sufficiently large constants that depend on s,t,d,k,ε, and d'. We will prove that

$$I(\mathcal{P},\mathcal{V}) \le \alpha_{1,d',k} m^{\frac{(d'-1)s}{d's-1}+\varepsilon} n^{\frac{d'(s-1)}{d's-1}} + \alpha_{2,d',k}(m+n). \tag{11.6}$$

The hidden constants in the $O(\cdot)$-notations throughout the proof may also depend on s,t,d,k, and ε. For brevity we write $O(\cdot)$ instead of $O_{s,t,d,k,\varepsilon}(\cdot)$.

For the induction basis, consider the case of $d' = 1$. Assuming that $m \ge s$, since the incidence graph contains no $K_{s,t}$ we get that at most $t-1$ varieties of $\mathcal{V}$ can contain U. By the Milnor-Thom theorem (Theorem 4.12), every other variety of $\mathcal{V}$ intersects U in $O(1)$ points and thus contributes $O(1)$ incidences. We conclude that in this case $I(\mathcal{P},\mathcal{V}) = O(m+n)$, and the claim holds when $\alpha_{1,1,k}$ and $\alpha_{2,1,k}$ are sufficiently large.

For the induction step, consider $d' > 1$ and assume that the theorem holds for smaller values of d'. We prove the induction step using a second induction on $m+n$. That is, we will prove (11.6) for a fixed $d' > 1$ by induction on the size of $\mathcal{P}$ and $\mathcal{V}$. For the induction basis, when both m and n are at most some constant the claim holds by taking sufficiently large $\alpha_{1,d',k}$ and $\alpha_{2,d',k}$.

We move to the induction step of induction on $m+n$. Since the incidence graph contains no copy of $K_{s,t}$, Theorem 7.1 implies $I(\mathcal{P},\mathcal{V}) = O(mn^{1-1/s})$. When $m = O(n^{1/s})$, this implies $I(\mathcal{P},\mathcal{V}) = O(n)$. We may thus assume that

$$n = O(m^s), \tag{11.7}$$

which in turn implies

$$n = n^{\frac{d'(s-1)}{d's-1}} n^{\frac{d'-1}{d's-1}} = O\left(m^{\frac{(d'-1)s}{d's-1}} n^{\frac{d'(s-1)}{d's-1}}\right). \tag{11.8}$$

Partitioning the space. Let $f \notin \mathbf{I}(U)$ be an r-partitioning polynomial of $\mathcal{P}$, for a sufficiently large constant r. The asymptotic relations between the various constants are

$$d, k, s, t, 2^{1/\varepsilon} \ll r \ll \alpha_{2,d',k} \ll \alpha_{1,d',k}.$$

By Theorem 11.4 we have $\deg f = O(r^{d/d'})$. Since $\mathcal{P} \subset U$, we define the cells of the partition to be the connected components of $U \setminus \mathbf{V}(f)$. By Theorem 4.11 with $W = \mathbf{V}(f)$, the number of cells is $c = O\left((r^{d/d'})^{d'}\right) = O\left(r^d\right)$. Denote the cells of the partition as $C_1, \ldots, C_c$. For each $j = 1, \ldots, c$, let $\mathcal{V}_j$ be the set of varieties of $\mathcal{V}$ that intersect C_j and set $\mathcal{P}_j = C_j \cap \mathcal{P}$. We also set $m_j = |\mathcal{P}_j|$, $m' = \sum_{j=1}^c m_j$, and $n_j = |\mathcal{V}_j|$. Note that $m_j \leq m/r^d$ for every $1 \leq j \leq c$. For any $W \in \mathcal{V}$, since $\dim(W \cap U) \leq d' - 1$, Theorem 4.11 implies that W intersects $O\left(r^{d(d'-1)/d'}\right)$ cells of $U \setminus \mathbf{V}(f)$ (recall that by assumption $U \not\subseteq W$). Therefore, $\sum_{j=1}^c n_j = O\left(nr^{d(d'-1)/d'}\right)$, and according to Hölder's inequality we have

$$\sum_{j=1}^c n_j^{\frac{d'(s-1)}{d's-1}} \leq \left(\sum_{j=1}^c n_j\right)^{\frac{d'(s-1)}{d's-1}} \left(\sum_{j=1}^c 1\right)^{\frac{d'-1}{d's-1}} = O\left(\left(nr^{d(d'-1)/d'}\right)^{\frac{d'(s-1)}{d's-1}} r^{\frac{d(d'-1)}{d's-1}}\right)$$

$$= O\left(n^{\frac{d'(s-1)}{d's-1}} r^{\frac{ds(d'-1)}{d's-1}}\right).$$

Combining the above with the hypothesis of the second induction implies

$$\sum_{j=1}^c I(\mathcal{P}_j, \mathcal{V}_j) \leq \sum_{j=1}^c \left(\alpha_{1,d',k} m_j^{\frac{(d'-1)s}{d's-1}+\varepsilon} n_j^{\frac{d'(s-1)}{d's-1}} + \alpha_{2,d',k}(m_j + n_j)\right)$$

$$\leq \alpha_{1,d',k} \frac{m^{\frac{(d'-1)s}{d's-1}+\varepsilon}}{r^{\frac{ds(d'-1)}{d's-1}+d\varepsilon}} \sum_{j=1}^c n_j^{\frac{d'(s-1)}{d's-1}} + \sum_{j=1}^c \alpha_{2,d',k}(m_j + n_j)$$

$$= O\left(\alpha_{1,d',k} \frac{m^{\frac{(d'-1)s}{d's-1}+\varepsilon} n^{\frac{d'(s-1)}{d's-1}}}{r^{d\varepsilon}}\right) + \alpha_{2,d',k}\left(m' + O\left(nr^{d(d'-1)/d'}\right)\right).$$

By applying (11.8) and taking $\alpha_{1,d',k}$ to be sufficiently large with respect to r and $\alpha_{2,d',k}$, we obtain

$$\sum_{j=1}^{c} I(\mathcal{P}_j, \mathcal{V}_j) = O\left(\alpha_{1,d',k} \frac{m^{\frac{(d'-1)s}{d's-1}+\varepsilon} n^{\frac{d'(s-1)}{d's-1}}}{r^{d\varepsilon}}\right) + \alpha_{2,d',k} m'.$$

When r is sufficiently large with respect to ε and to the constant hidden in the $O(\cdot)$-notation, we have

$$\sum_{j=1}^{c} I(\mathcal{P}_j, \mathcal{V}_j) \leq \frac{\alpha_{1,d',k}}{2} m^{\frac{(d'-1)s}{d's-1}+\varepsilon} n^{\frac{d'(s-1)}{d's-1}} + \alpha_{2,d',k} m'. \tag{11.9}$$

Incidences on the partition. Let $U_0 = U \cap \mathbf{V}(f)$, $\mathcal{P}_0 = \mathcal{P} \cap U_0$, and $m_0 = |\mathcal{P}_0| = m - m'$. Let $\mathcal{V}_0$ denote the set of varieties of $\mathcal{V}$ that are contained in $\mathbf{V}(f)$. Set $\mathcal{V}' = \mathcal{V} \setminus \mathcal{V}_0$, $n_0 = |\mathcal{V}_0|$, and $n' = |\mathcal{V}'| = n - n_0$.

It remains to bound incidences with the points of $\mathcal{P}_0$. We partition these incidences into two types:

- Let I_1 be the set of incidences $(p, W) \in \mathcal{P}_0 \times \mathcal{V}$ such that some irreducible component of U_0 contains p and is contained in W.
- Let I_2 be the set of incidences $(p, W) \in \mathcal{P}_0 \times \mathcal{V}$ such that no irreducible component of U_0 contains p and is contained in W.

Note that $I(\mathcal{P}_0, \mathcal{V}) = |I_1| + |I_2|$.

We first derive an upper bound for $|I_1|$. Since U is an irreducible variety and $U \not\subseteq \mathbf{I}(f)$, we get that U_0 is a variety of dimension at most $d' - 1$ and of degree $k_0 = O(r^{d/d'})$. By Lemma 4.9, the number of irreducible components of U_0 is $O_r(1)$. The degree of each such component is at most k_0.

Consider an irreducible component A of U_0. If A contains at most $s - 1$ points of $\mathcal{P}_0$, then these points contribute at most $(s-1)n$ incidences to I_1. If A contains at least s points of $\mathcal{P}_0$ then at most $t - 1$ varieties of $\mathcal{V}$ contain A. In this case, the points in A contribute at most $(t-1)m_0$ incidences to I_1. By summing these bounds over every irreducible component of U_0, choosing sufficiently large $\alpha_{1,d',k}$ and $\alpha_{2,d',k}$, and recalling (11.8), we have

$$|I_1| = O_r\left(n + m_0\right) < \frac{\alpha_{1,d',k}}{4} m^{\frac{s(d'-1)}{d's-1}} n^{\frac{d'(s-1)}{d's-1}} + \frac{\alpha_{2,d',k}}{2} m_0. \tag{11.10}$$

We next derive an upper bound for $|I_2|$. Because U_0 is of dimension at most $d' - 1$, we can apply the induction hypothesis with each irreducible component A of

U_0. Specifically, we apply the hypothesis in A with the point set $\mathcal{P}_0 \cap A$ and the set of varieties of $\mathcal{V}$ that do not contain A. Since U_0 has $O_r(1)$ irreducible components and each is of degree at most k_0, we get

$$|I_2| = O_r\left(\alpha_{1,d'-1,k_0} m_0^{\frac{s(d'-2)}{(d'-1)s-1}+\varepsilon} n^{\frac{(d'-1)(s-1)}{(d'-1)s-1}} + \alpha_{2,d'-1,k_0}(m_0+n)\right). \tag{11.11}$$

By (11.7) we have that

$$m_0^{\frac{s(d'-2)}{(d'-1)s-1}+\varepsilon} n^{\frac{(d'-1)(s-1)}{(d'-1)s-1}} \leq m^{\frac{s(d'-2)}{(d'-1)s-1}+\varepsilon} n^{\frac{d'(s-1)}{d's-1}} n^{\frac{s-1}{(d's-s-1)(d's-1)}} = O\left(m^{\frac{s(d'-1)}{d's-1}+\varepsilon} n^{\frac{d'(s-1)}{d's-1}}\right).$$

By combining this with (11.8) and (11.11), and taking $\alpha_{1,d',k}$ and $\alpha_{2,d',k}$ to be sufficiently large with respect to $\alpha_{1,d'-1,k_0}$ and $\alpha_{2,d'-1,k_0}$, we obtain

$$|I_2| \leq \frac{\alpha_{1,d',k}}{4} m^{\frac{s(d'-1)}{d's-1}+\varepsilon} n^{\frac{d'(s-1)}{d's-1}} + \frac{\alpha_{2,d',k}}{2} m_0. \tag{11.12}$$

Combining (11.9), (11.10), and (11.12) completes the induction step and the proof of the theorem. $\qquad\square$

11.4   Exercises

Problem 11.1. (a) Let m be a positive integer and let $\ell \subset \mathbb{R}^2$ be a line. Find the dimension of the vector space $\mathbf{I}(\ell)_{\leq m}$ (as a function of m).
(b) Find the Hilbert polynomial and the regularity of the ideal $J = \langle x^3, x^2 y \rangle$.

Problem 11.2. Let $\mathcal{P}$ be a set of m points and let $\mathcal{V}$ be a set of n varieties of degree at most k and dimension at most $d' < d$, both in $\mathbb{R}^d$. The incidences graph of $\mathcal{P} \times \mathcal{V}$ does not contain a copy of $K_{s,t}$. Prove that for every $\varepsilon > 0$, we have

$$I(\mathcal{P}, \mathcal{V}) = O_{k,s,t,d,\varepsilon}\left(m^{\frac{d's}{(d'+1)s-1}+\varepsilon} n^{\frac{(d'+1)(s-1)}{(d'+1)s-1}} + m + n\right).$$

(Hint: There is no need to use an involved analysis — there is a short proof.)

Problem 11.3. Let $\mathcal{P}$ be a set of m points and let Γ be a set of n circles in $\mathbb{R}^2$. While we learned in Chapter 3 how to derive an upper bound for $I(\mathcal{P}, \Gamma)$, in this problem you are asked to study a dual proof.

Given a circle $\gamma = \mathbf{I}\left((x-a)^2 + (y-b)^2 - r^2\right)$, define the dual point of γ as $\gamma^* = (a, b, r^2) \in \mathbb{R}^3$. Consider the point set $\Gamma^* = \{\gamma^* : \gamma \in \Gamma\}$. Define a dual set

$\mathcal{P}^*$ of varieties and derive a bound for $I(\mathcal{P}, \Gamma)$ by obtaining a bound for the dual incidence problem.

Problem 11.4. Change the proof of Theorem 11.5 so that it would show the dependency of the bound in t. In particular, prove that

$$I(\mathcal{P}, \mathcal{V}) = O_{k,s,d,\varepsilon}\left(m^{\frac{(d'-1)s}{d's-1}+\varepsilon} n^{\frac{d'(s-1)}{d's-1}} t^{\frac{d'-1}{d's-1}} + mt + n \right).$$

Problem 11.5. Let $\mathcal{P}$ be a set of m points in $\mathbb{R}^3$. Prove that there exists a variety of dimension one and degree $O(\sqrt{m})$ that contains $\mathcal{P}$ (hint: Recall Lemma 5.4 and use (11.5) as in the proof of Lemma 11.3).

11.5   Open problems

In Section 11.1 we mentioned conjectured bounds for point-variety incidences in $\mathbb{R}^d$, depending on the dimension of the varieties. We now discussion these conjectured bounds in more detail. Let $\mathcal{P}$ be a set of m points and let $\mathcal{V}$ be a set of n varieties of degree at most k and dimension d', both in $\mathbb{R}^d$. Assume that the incidence graph of $\mathcal{P} \times \mathcal{V}$ contains no copy of $K_{s,t}$. To bound the number of incidences in $\mathcal{P} \times \mathcal{V}$, we use the partitioning polynomial technique as studied in Chapter 3 (that is, a partitioning not of a constant degree). Specifically, we use an r-partitioning polynomial to partition $\mathbb{R}^d$ into $O(r^d)$ cells and then apply the weak combinatorial bound of Lemma 7.1 separately in each cell. A simple calculation shows that setting $r = m^{\frac{s}{ds-d+d'}} n^{\frac{-1}{ds-d+d'}}$ implies that the total number of incidences in the cells is

$$O_{d,k,s,t}\left(m^{\frac{sd'}{ds-d+d'}} n^{\frac{ds-d}{ds-d+d'}} + m + n \right). \tag{11.13}$$

As we have seen, when dealing with incidences in dimension $d \geq 3$ the most difficult part of the analysis is handling the incidences that are on the partition. Thus, the above bound for the number of incidences inside the cells might be meaningless. Nonetheless, this bound does match the bounds of the three cases that were already established (up to the extra ε in the exponent):

- Theorem 11.2 matches the bound of (11.13) in the case where $d' = d - 1$.
- Theorem 7.5 matches the bound of (11.13) in the case where $d' = d/2$. This requires an extra assumption regarding the tangents of the varieties.
- The simple extension of Theorem 11.1 to $\mathbb{R}^d$ matches the bound of (11.13) in the case where $d' = 1$. This requires an extra assumption regarding the number of varieties of $\mathcal{V}$ that can be contained in a low-degree variety.

It seems reasonable to make the following conjecture.

Conjecture 11.6. *Let $\mathcal{P}$ be a set of m points and let $\mathcal{V}$ be a set of n varieties of degree at most k and dimension d', both in $\mathbb{R}^d$. Assume that the incidence graph contains no copy of $K_{s,t}$ and that the varieties of $\mathcal{V}$ satisfy some reasonable conditions. Then for any $\varepsilon > 0$ we have*

$$I(\mathcal{P}, \mathcal{V}) = O_{d,k,s,t,\varepsilon}\left(m^{\frac{sd'}{ds-d+d'}} n^{\frac{ds-d}{ds-d+d'}} + m + n\right).$$

The statement "the varieties of $\mathcal{V}$ satisfy some reasonable conditions" is not well defined, and it is not clear what the "right" conditions are. Possible conditions might be a bound on the number of elements of $\mathcal{V}$ in a low degree variety and/or a condition about the tangent spaces as in Theorem 7.5. While one can make a more rigorous conjecture regarding what these conditions should be, we do not state such a conjecture here.

The distinct distances problem in $\mathbb{R}^d$ can be reduced to an incidence problem with $(d-1)$-flats in $\mathbb{R}^{2d-1}$ (see [4]). This problem and others suggest that the case of $d' = (d-1)/2$ (for odd d) is a main open case of Conjecture 11.6.

Theorem 11.2 is known to be tight up to an ε in the exponent when $s = 2$ and $d \geq 4$, for many types of hypersurfaces (see [84]). When $s \geq 3$, it is not clear whether the bound of Theorem 11.2 is close to being tight. Recall from Section 3.6 that the planar variant of Theorem 11.2 (Theorem 3.3) is tight for some types of curves when $s = 2$, but that stronger bounds are known when $s \geq 3$.

Open Problem 11.1. Find whether the bound of Theorem 11.2 is tight when $s \geq 3$ and $d \geq 3$. That is, either find a construction that matches the bound of the theorem or derive an improved upper bound.

We already know that the ε in Theorem 11.2 can be removed when $d = 2$ (this is the planar bound from Section 3.6). Zahl [107] showed that the ε can be removed from the incidence bound when $d = 3$ and assuming that the intersection of every three varieties is a finite point set. Basu and Sombra [7] extended this to the case of $d = 4$, when assuming that the intersection of every four varieties is a finite point set.

Open Problem 11.2. Find whether the ε in the exponent of Theorem 11.2 can be removed when $d \geq 5$, possibly under additional (reasonable) assumptions.

Chapter 12

Applications in $\mathbb{R}^d$

> *"The older I get, the more I believe that at the bottom of most deep mathematical problems there is a combinatorial problem." / Israel Gelfand.*

In Chapter 11 we derived bounds for point-variety incidences in $\mathbb{R}^d$. In particular, Theorem 11.2 is a general point-variety incidence bound in $\mathbb{R}^d$. In the current chapter we study a couple of applications for that theorem. These applications do not require reading any part of Chapter 11 beyond the statement of Theorem 11.2.

The first application comes from Discrete Geometry, and is a distinct distances problem. The second application comes from a discrete Fourier restriction problem in Harmonic Analysis (for simplicity, we only discuss the combinatorial part of this problem).

12.1   Distinct distances with local properties

Our first application concerns a family of distinct distances problems that were posed by Erdős [36]. Let $\phi(n, k, \ell)$ denote the minimum number of distinct distances that are determined by a set $\mathcal{P}$ of n points in $\mathbb{R}^2$, such that any k points of $\mathcal{P}$ determine at least ℓ distinct distances. That is, we wish to show that a local property for small subsets of points can be used to obtain a global property of the entire point set.

Figure 12.1: A set of eight points that determine few distinct distances, with every three points determining three distinct distances.

138

For example, $\phi(n,3,3)$ is the minimum number of distinct distances that are determined by a set of n points that do not span any isosceles triangles (including equilateral triangles and degenerate triangles with three collinear vertices). Let $\mathcal{P}$ be a point set that does not span any isosceles triangles and let $p \in \mathcal{P}$. Since there are no isosceles triangles, p cannot be at the same distance from two points of $\mathcal{P} \setminus \{p\}$. This immediately implies $\phi(n,3,3) \geq n-1$. Behrend [10] proved that there exists a set A of positive integers $a_1 < a_2 < \cdots < a_n$ such that no three elements of A determine an arithmetic progression and $a_n < n2^{O(\sqrt{\log n})}$. Since no three points form an arithmetic progression, the point set $\mathcal{P}_1 = \{(a_1,0),(a_2,0),\ldots,(a_n,0)\}$ does not span any isosceles triangles (see Figure 12.1 for an illustration). Since $\mathcal{P}_1 \subset \mathcal{P}_2 = \{(1,0),(2,0),\ldots,(a_n,0)\}$ and $D(\mathcal{P}_2) < n2^{O(\sqrt{\log n})}$, we get that $D(\mathcal{P}_1) < n2^{O(\sqrt{\log n})}$. That is, $\phi(n,3,3) < n2^{O(\sqrt{\log n})}$. Closing the remaining gap between this bound and $\phi(n,3,3) = \Omega(n)$ is an open problem.

For a constant $k \geq 4$, consider the case of $\phi\left(n,k,\binom{k}{2}-k+3\right)$. Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^2$ such that every k points of $\mathcal{P}$ span at least $\binom{k}{2}-k+3$ distinct distances. Let $p \in \mathcal{P}$ and assume that there exists a distance δ such that p is at distance δ from at least $k-1$ points of $\mathcal{P} \setminus \{p\}$. Let $\mathcal{P}' \subset \mathcal{P}$ be a set of $k-1$ points that are at a distance of δ from p. Then $\mathcal{P}' \cup \{p\}$ is a set of k points that determine at most $\binom{k}{2}-k+2$ distinct distances. This contradicts the assumption on $\mathcal{P}$, so p can span any given distance with at most $k-2$ points of $\mathcal{P}$. Since k is a constant, this immediately implies that

$$\phi\left(n,k,\binom{k}{2}-k+3\right) \geq \frac{n}{k-2} = \Omega(n).$$

We now derive a stronger bound, which is a variant of a result by Fox, Pach, and Suk [43].

Theorem 12.1. *For any $k \geq 6$ and $\varepsilon > 0$,*

$$\phi\left(n,k,\binom{k}{2}-k+5\right) = \Omega_{k,\varepsilon}\left(n^{8/7-\varepsilon}\right).$$

Proof. Let $\mathcal{P}$ be a set of n points in $\mathbb{R}^2$ such that every k points of $\mathcal{P}$ span at least $\binom{k}{2}-k+5$ distinct distances. For points $a,p \in \mathbb{R}^2$, we denote by $|ap|$ the distance between a and p. Let

$$Q = \left\{(a,p,b,q) \in \mathcal{P}^4 : |ap| = |bq| > 0\right\}.$$

The quadruples are ordered, so (a,p,b,q) and (p,a,b,q) are in Q as distinct quadruples.

Let x denote the number of distinct distances that are determined by $\mathcal{P}$, and denote these distances as $\delta_1, \ldots, \delta_x$. For $1 \le j \le x$, let $E_j = \{(a,b) \in \mathcal{P}^2 : |ab| = \delta_j\}$. Since every ordered pair of $\mathcal{P}^2$ appears in exactly one E_j, we have $\sum_{j=1}^{x} |E_j| = n^2$. Note that the number of quadruples in Q that satisfy $|ap| = |bq| = \delta_j$ is exactly $|E_j|^2$. By the Cauchy–Schwarz inequality, we have

$$|Q| = \sum_{j=1}^{x} |E_j|^2 \ge \frac{\left(\sum_{j=1}^{x} |E_j|\right)^2}{x} = \frac{n^4}{x}. \tag{12.1}$$

Defining Q and using Cauchy–Schwarz to obtain a lower bound for $|Q|$ is a standard approach for distinct distances problems and other related problems (see Chapters 8 and 10). To complete the proof, it remains to derive an upper bound for $|Q|$. We do this by presenting a reduction to an incidence problem.

For two points $a, b \in \mathcal{P}$ we define the point $v_{a,b} = (a_x, a_y, b_x, b_y) \in \mathbb{R}^4$, and consider the set of points

$$\mathcal{P}_4 = \{v_{a,b} : a, b \in \mathcal{P} \text{ and } a \ne b\} \subset \mathbb{R}^4.$$

For two points $p, q \in \mathcal{P}$ we define the variety

$$S_{p,q} = \mathbf{V}\left((x_1 - p_x)^2 + (x_2 - p_y)^2 - (x_3 - q_x)^2 - (x_4 - q_y)^2\right) \subset \mathbb{R}^4.$$

Finally, we consider the set of varieties

$$\mathcal{V} = \{S_{p,q} : p, q \in \mathcal{P} \text{ and } p \ne q\}.$$

Note that a point $v_{a,b}$ is incident to a variety $S_{p,q}$ if and only if $|ap| = |pq|$. This gives a bijection between the incidences of $\mathcal{P}_4 \times \mathcal{V}$ and the quadruples of Q. That is, to derive an upper bound on $|Q|$ it suffices to obtain an upper bound for $I(\mathcal{P}_4, \mathcal{V})$.

Assume that there is a copy of $K_{2,k-4}$ in the incidence graph of $\mathcal{P}_4 \times \mathcal{V}$. This means that there is a set of k points $\{a_1, b_1, a_2, b_2, p_1, q_1, \ldots, p_{(k-4)/2}, q_{(k-4)/2}\} \subset \mathcal{P}$ such that $|a_j p_\ell| = |b_j q_\ell|$ for every $1 \le j \le 2$ and $1 \le \ell \le (k-4)/2$. This is a set of k points that determines at most $\binom{k}{2} - k + 4$ distinct distances, which contradicts the assumption on $\mathcal{P}$. This contradiction implies that the incidence graph of $\mathcal{P}_4 \times \mathcal{V}$ contains no copy of $K_{2,k-4}$.

Note that $|\mathcal{P}_4| = \Theta(n^2)$ and that $|\mathcal{V}| = \Theta(n^2)$. By applying Theorem 11.2 on $\mathcal{P}_4$ and $\mathcal{V}$ with $d = 4$, $s = 2$, and $t = k - 4$, we obtain

$$I(\mathcal{P}_4, \mathcal{V}) = O_{k,\varepsilon}\left((n^2)^{6/7+\varepsilon}(n^2)^{4/7} + n^2\right) = O_{k,\varepsilon}\left(n^{20/7+\varepsilon}\right).$$

Combining this with (12.1) gives $x = \Omega\left(n^{8/7-\varepsilon}\right)$, which completes the proof. $\qquad\square$

12.2 Additive energy on a hypersphere

As stated in the beginning of the chapter, our next application comes from a discrete Fourier restriction problem in Harmonic Analysis. In particular, it is taken from a work of Bourgain and Demeter [15]. Since restriction problems are outside the scope of this book, here we remove the analytical context and only present the underlying combinatorial result. We present the problem in $\mathbb{R}^4$, which is the lowest dimensional space in which the following approach holds (a lower dimensional variant of the problem is mentioned in Section 12.4).

Given a finite set $\mathcal{P} \subset \mathbb{R}^d$, the *additive energy* of $\mathcal{P}$ is

$$E(\mathcal{P}) = \left| \left\{ (a, b, p, q) \in \mathcal{P}^4 \ : \ a + b = p + q \right\} \right|.$$

The additive energy of a set is a main object in Additive Combinatorics. It is strongly related to the additive structure of $\mathcal{P}$, and in particular to the sum set $\mathcal{P} + \mathcal{P} = \{a + b : a, b \in \mathcal{P}\}$. We will study additive energy in more detail in Section 13.5 (see also [99]).

We denote the coordinates of $\mathbb{R}^4$ as (x_1, x_2, x_3, x_4) and for a point $p \in \mathbb{R}^4$ write $p = (p_1, p_2, p_3, p_4)$. For a positive integer n, we consider the hypersphere

$$S_n = \mathbf{V}\left(x_1^2 + x_2^2 + x_3^2 + x_4^2 - n^2\right) \subset \mathbb{R}^4.$$

In other words, S_n is a hypersphere centered at the origin and of radius n. We are interested in sets of points of S_n that have only integer coordinates.

Theorem 12.2. *Let $\mathcal{P} \subset S_n \cap \mathbb{Z}^4$. Then for every $\varepsilon > 0$ we have*

$$E(\mathcal{P}) = O\left(|\mathcal{P}|^{7/3} n^\varepsilon\right).$$

To prove Theorem 12.2, we first need to know how large can $S_n \cap \mathbb{Z}^4$ be. This is equivalent to asking how many representations n^2 has as a sum of four squares. Such number theoretic problems are mostly solved, as the following result states (see for example [46]. For part (d) see [84]).

Theorem 12.3. *There exists a constant c such that the following hold for any positive integer n and for every $\varepsilon > 0$.*
(a) Any circle in $\mathbb{R}^2$ of radius n and center in $\mathbb{Z}^2$ contains $O(n^{c/\log \log n})$ points of $\mathbb{Z}^2$.
(b) Any sphere in $\mathbb{R}^3$ of radius n and center in $\mathbb{Z}^3$ contains $O(n^{1+\varepsilon})$ points of $\mathbb{Z}^3$.
(c) For $d \geq 4$, any hypersphere in $\mathbb{R}^d$ of radius n and center in $\mathbb{Z}^d$ contains $O(n^{d-2})$ points of $\mathbb{Z}^d$.
(d) Let γ be a circle in $\mathbb{R}^4$ that is the intersection of two hyperplanes with the hypersphere centered at the origin and of radius n. The hyperplanes are defined by

linear equations with integer coefficients whose absolute values have size $O(n)$. Then γ contains $O(n^{c/\log\log n})$ points of $\mathbb{Z}^4$.

We will also need a variant of Theorem 11.2 that shows the exact dependency of the bound in t. Proving this variant requires only minor changes to the proof that is presented in Section 11.3 (Problem 11.4 asks to prove this).

Theorem 12.4. *Let $\mathcal{P}$ be a set of m points and let $\mathcal{V}$ be a set of n varieties of degree at most k, both in $\mathbb{R}^d$. Assume that the incidences graph of $\mathcal{P} \times \mathcal{V}$ does not contain a copy of $K_{s,t}$. Then for every $\varepsilon > 0$, we have*

$$I(\mathcal{P}, \mathcal{V}) = O_{k,s,d,\varepsilon}\left(m^{\frac{(d-1)s}{ds-1}+\varepsilon} n^{\frac{d(s-1)}{ds-1}} t^{\frac{d-1}{ds-1}} + mt + n\right).$$

We are now ready to prove Theorem 12.2. In the proof of Theorem 12.1 from the preceding section, the reduction to an incidence problem was relatively simple. In the following proof we have to work harder to get to an incidence problem.

Proof of Theorem 12.2. Set $m = |\mathcal{P}|$. For every $v \in \mathbb{R}^4$ let m_v denote the number of pairs $(p, q) \in \mathcal{P}^2$ such that $v = p + q$. Given a fixed $v \in \mathbb{R}^4$, for every $p \in \mathcal{P}$ there is at most one $q \in \mathcal{P}$ such that $p + q = v$, so $m_v \le m$ (the bound is m rather than $m/2$ since the order of the pair (p, q) matters). Note that the number of quadruples (a, b, p, q) that satisfy $a + b = p + q = v$ is exactly m_v^2. For an integer $j \ge 1$, let $k_j = |\{v \in \mathbb{R}^4 : m_v \ge j\}|$. That is, k_j is the number of points in $\mathbb{R}^4$ that can be written as a sum of two points of $\mathcal{P}$ in at least j distinct ways. A dyadic pigeonholing argument gives[1]

$$E(\mathcal{P}) = \sum_{v \in \mathbb{R}^4} m_v^2 = \sum_{j=1}^{1+\log m} \sum_{\substack{v \in \mathbb{R}^4 \\ 2^{j-1} \le m_v < 2^j}} m_v^2 < \sum_{j=1}^{1+\log m} 2^{2j} k_{2^{j-1}}. \tag{12.2}$$

To derive an upper bound on $E(\mathcal{P})$, it remains to derive an upper bound on k_j. Since there are m^2 ordered pairs $(p, q) \in \mathcal{P}^2$ and each contributes to m_v for exactly one $v \in \mathbb{R}^4$, we have that $\sum_{v \in \mathbb{R}^4} m_v \le m^2$. Since every v that contributes to k_j has at least j corresponding pairs in $\mathcal{P}^2$, we get the bound

$$k_j \le \frac{m^2}{j}. \tag{12.3}$$

Denote the origin of $\mathbb{R}^4$ as o. The number of pairs $(p, q) \in \mathcal{P}$ that satisfy $p + q = o$ is at most m, these pairs contribute to $E(\mathcal{P})$ at most m^2 quadruples. We may thus assume that $v \ne o$ for the remainder of the proof.

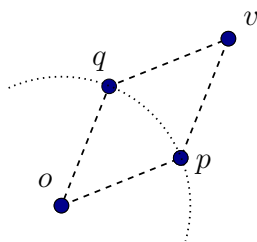[1] All of the logarithms in this proof are with base 2.

Figure 12.2: The quadrilateral $vpqo$ is a rhombus.

Consider $p, q \in \mathcal{P}$ and let $v = p + q$. Note that v is in the two-dimensional plane spanned by p, q, and o. Since $p, q \in S_n$ we have that $|op| = |oq| = n$, which in turn implies that the quadrilateral $vpqo$ is a rhombus of side length n. Figure 12.2 depicts such a configuration.

Let S_v be the hypersphere in $\mathbb{R}^4$ centered at v and of radius n. By the previous paragraph, points $p, q \in \mathcal{P}$ that satisfy $p + q = v$ are incident to S_v. That is, every two points $p, q \in \mathcal{P}$ that satisfy $p + q = v$ must be contained in the two-dimensional sphere $S_n \cap S_v$ centered at $v/2$ (in the degenerate case of $p = q$, the intersection $S_n \cap S_v$ is a single point). Let H_v be the unique hyperplane in $\mathbb{R}^4$ that contains the sphere $S_n \cap S_v$. By the above, if two points $p, q \in \mathcal{P}$ satisfy $p + q = v$ then p and q are incident to H_v. This implies that $m_v \leq |\mathcal{P} \cap H_v|$.

By the conclusion of the previous paragraph, to obtain an upper bound for k_j it suffices to bound the number of hyperplanes in $\mathbb{R}^4$ that contain at least j points of $\mathcal{P}$. This allows us to use incidences to complete the proof. We already used incidences to study a similar problem in Lemma 1.15.

An incidence argument. The sphere $S_n \cap S_v$ is also the set of points that are at a distance of n from both o and v. Thus, the containing hyperplane H_v is the perpendicular bisector of o and v (that is, the set of points in $\mathbb{R}^4$ that are at the same distance from o and from v). In particular, H_v is incident to the point $v/2$ and is orthogonal to v. That is,

$$H_v = \mathbf{V}(v_1(2x_1 - v_1) + v_2(2x_2 - v_2) + v_3(2x_3 - v_3) + v_4(2x_4 - v_4)).$$

Given distinct $u, v \in \mathbb{R}^4 \setminus \{0\}$, the intersection $H_v \cap H_u \cap S_n$ is either a circle or a set of at most two points. Moreover, if the absolute value of any coordinate of v is larger than $2n$ then $m_v = 0$ (and symmetrically for u). We may thus apply Theorem 12.3(d), to obtain that $H_v \cap H_u \cap S_n$ contains $O(n^{c/\log\log n})$ points of $\mathbb{Z}^4$. We will use the weaker bound $O(n^{\varepsilon'})$ rather than $O(n^{c/\log\log n})$, where $\varepsilon' = \varepsilon/100$.

Note that no plane H_v is incident to the origin, since H_v is the perpendicular bisector of the origin and v. Consider an integer $j = \Omega_{\varepsilon'}(1)$. Let Π_j be the set of

hyperplanes in $\mathbb{R}^4$ that contain at least j points of $\mathcal{P}$ and are not incident to the origin. By the above, $k_j \leq |\Pi_j|$. We will derive an upper bound for $|\Pi_j|$ by using a point-hyperplane incidence bound in $\mathbb{R}^4$. By the preceding paragraph, the incidence graph of $\mathcal{P} \times \Pi_j$ contains no copy of $K_{t,2}$ where $t = O(n^{\varepsilon'})$.

To replace the "no $K_{t,2}$" condition with "no $K_{2,t}$", we move to a dual space as follows (see the proof of Lemma 13.1 for a similar argument). For a point $p \in \mathbb{R}^4$ we define the dual hyperplane as $p^* = \mathbf{V}(p_1 x_1 + p_2 x_2 + p_3 x_3 + p_4 x_4 - 1) \subset \mathbb{R}^4$. Consider a hyperplane $H \in \Pi_j$, and recall that H is not incident to the origin. Thus, we can write $H = \mathbf{V}(a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 - 1)$ for some $a_1, a_2, a_3, a_4 \in \mathbb{R}$. We define the dual point of H as $H^* = (a_1, a_2, a_3, a_4) \in \mathbb{R}^4$.

Consider the set of planes

$$\mathcal{P}^* = \{p^* \ : \ p \in \mathcal{P}\},$$

and the set of points

$$\Pi_j^* = \{H^* \ : \ H \in \Pi_j\}.$$

Let $p \in \mathcal{P}$ and $H \in \Pi_j$. We have that $p \in H$ if and only if $H^* \in p^*$, since both are equivalent to $a_1 p_1 + a_2 p_2 + a_3 p_3 + a_4 p_4 = 1$. This implies that $I(\mathcal{P}, \Pi_j) = I(\Pi_j^*, \mathcal{P}^*)$. In addition, since the incidence graph of $\mathcal{P} \times \Pi_j$ does not contain $K_{t,2}$, the incidence graph of $\Pi_j^* \times \mathcal{P}^*$ does not contain a copy of $K_{2,t}$. Applying Theorem 12.4 on Π_j^* and $\mathcal{P}^*$ with $t = O(n^{\varepsilon'})$ implies

$$I(\Pi_j^*, \mathcal{P}^*) = O_{\varepsilon'}\left(|\Pi_j|^{6/7+\varepsilon'} m^{4/7} n^{\varepsilon'} + |\Pi_j| n^{\varepsilon'} + m\right).$$

Since each hyperplane of Π_j is incident to at least j points of $\mathcal{P}$, we have that

$$I(\Pi_j^*, \mathcal{P}^*) = I(\mathcal{P}, \Pi_j) \geq j|\Pi_j|.$$

Combining the two above bounds for $I(\Pi_j^*, \mathcal{P}^*)$ leads to

$$j|\Pi_j| = O_{\varepsilon'}\left(|\Pi_j|^{6/7+\varepsilon'} m^{4/7} n^{\varepsilon'} + |\Pi_j| n^{\varepsilon'} + m\right). \tag{12.4}$$

Since $j = \Omega_{\varepsilon'}(1)$, the right-hand side of (12.4) cannot be dominated by the term $|\Pi_j| n^{\varepsilon'}$. Removing the term $|\Pi_j| n^{\varepsilon'}$ from (12.4) and rearranging gives

$$k_j \leq |\Pi_j| = O_{\varepsilon'}\left(\frac{m^{\frac{4}{1-7\varepsilon'}} n^{\frac{7\varepsilon'}{1-7\varepsilon'}}}{j^{\frac{7}{1-7\varepsilon'}}} + \frac{m}{j}\right) = O_{\varepsilon'}\left(\frac{m^{\frac{4}{1-7\varepsilon'}} n^{\frac{7\varepsilon'}{1-7\varepsilon'}}}{j^7} + \frac{m}{j}\right).$$

The statement of the theorem is trivial when $\varepsilon \geq 1$. We may thus assume that $\varepsilon < 1$, which implies $\varepsilon' < 1/100$. For such ε' we have $\frac{1}{1-7\varepsilon'} \leq (1 + 10\varepsilon')$, so

$$k_j = O_{\varepsilon'}\left(\frac{m^{4+40\varepsilon'}n^{8\varepsilon'}}{j^7} + \frac{m}{j}\right). \tag{12.5}$$

Completing the proof. We now use (12.3) and (12.5) to bound the elements $k_{2^{j-1}}$ in (12.2). In particular, when $j \leq \log m^{1/3}$ we apply (12.2) and for larger values of j we apply (12.5). This gives

$$E(\mathcal{P}) = O_{\varepsilon'}\left(\sum_{j=1}^{\log m^{1/3}} 2^j m^2 + \sum_{j=1+\log m^{1/3}}^{1+\log m}\left(\frac{m^{4+40\varepsilon'}n^{8\varepsilon'}}{2^{5j}} + 2^j m\right)\right).$$

$$= O_{\varepsilon'}\left(m^{7/3} + m^{7/3+40\varepsilon'}n^{8\varepsilon'} + m^2\right) = O_{\varepsilon'}\left(m^{7/3+40\varepsilon'}n^{8\varepsilon'}\right).$$

Recall that $m = \mathcal{P} \subset S_n \cap \mathbb{Z}^4$. By Theorem 12.3(c), we have that $m = O(n^2)$. This implies that $m^{40\varepsilon'}n^{8\varepsilon'} = O(n^{88\varepsilon'}) = O(n^\varepsilon)$. We get that $E(\mathcal{P}) = O_\varepsilon\left(m^{7/3}n^\varepsilon\right)$, as asserted. $\qquad\square$

12.3 Exercises

Problem 12.1. Find the exact asymptotic value of $\phi\left(n, k, \binom{k}{2} - \lfloor k/2 \rfloor + 2\right)$ for every $k \geq 4$.

Problem 12.2. Prove that $\phi\left(n, k, \binom{k}{2} - \lfloor 2k/3 \rfloor + 3\right) = \Omega(n^{3/2})$ for every $k \geq 6$. One way to do this is to assume that there exists a distance δ that is spanned by $\Omega(n^{1/2})$ pairs of points (why are we allowed to assume that?). Then consider the subset of points that are at distance δ from at least one other point.

Problem 12.3. Let S be a sphere in $\mathbb{R}^3$ centered at the origin. Prove that for any finite point set $\mathcal{P} \subset S$ and $\varepsilon > 0$, we have that $E(\mathcal{P}) = O(|\mathcal{P}|^{20/9+\varepsilon})$ (note that the points of ε are not required to have integer coordinates).

Hint: Theorem 3.9 with $s = 3$ gives a point-circle incidence bound in $\mathbb{R}^2$. Why is this bound still valid in $\mathbb{R}^3$?

12.4 Open problems

Both of the problems that were studied in this chapter are still open. Section 12.1 introduced a family of problems concerning distinct distances for point sets with local

properties. In other words, studying the asymptotic value of $\phi(n, k, \ell)$ for various values of k and ℓ. The following seem to be considered as the main open cases.

Open Problem 12.1.
(a) Find the asymptotic size of $\phi(n, 4, 5)$.
(b) Find the asymptotic size of $\phi(n, 5, 9)$.

No non-trivial bounds is known for either part of Open Problem 12.1. Specifically, in both cases we have the trivial upper bound $O(n^2)$ and a lower bound of $\Omega(n)$.

The more general problem that was studied in Theorem 12.1 is also wide open. While the theorem provides a lower bound of $\Omega\left(n^{8/7-\varepsilon}\right)$, this is still far from the upper bound $O(n^2)$ (no non-trivial upper bound is known).

Open Problem 12.2. Find the asymptotic size of $\phi\left(n, k, \binom{k}{2} - k + c\right)$ where k is sufficiently large and c is constant.

More information about this family of distinct distances problems can be found in [85].

We now move to the energy problem that was presented in Section 12.2. For any finite set $\mathcal{P} \subset \mathbb{R}^4$ we have the trivial lower bound $E(\mathcal{P}) = \Omega(|\mathcal{P}|^2)$, obtained from quadruples $(a, b, p, q) \in \mathcal{P}^4$ with $a = p$ and $b = q$. This bound is much smaller than the bound $E(\mathcal{P}) = O\left(|\mathcal{P}|^{7/3} n^\varepsilon\right)$ of Theorem 12.2. Curiously, Bourgain and Demeter [15] showed that when replacing the sphere S_n with a truncated paraboloid $\mathbf{V}(x_1^2 + x_2^2 + x_3^2 - x_4)$, there are point sets $\mathcal{P}$ that satisfy $E(\mathcal{P}) = \Omega(|\mathcal{P}|^{7/3})$.

Open Problem 12.3. Let $\mathcal{P} \subset S_n \cap \mathbb{Z}^4$. Find the maximum asymptotic size that $E(\mathcal{P})$ can have.

Another interesting variant of the problem is in $\mathbb{R}^3$.

Open Problem 12.4. Let S be a sphere in $\mathbb{R}^3$ centered at the origin and let $\mathcal{P} \subset S$ be a finite set (not necessarily in $\mathbb{Z}^3$). Find the maximum asymptotic size that $E(\mathcal{P})$ can have.

Problem 12.3 gives the current best bound for the three-dimensional problem. When replacing the sphere with a paraboloid, Demeter derived the bound $E(\mathcal{P}) = O(|\mathcal{P}|^{2+\varepsilon})$ (for any $\varepsilon > 0$). More information about this family of problems can be found in [25].

Chapter 13

Incidences in Spaces Over Finite Fields

In Chapter 6 we started the study of incidence-related problems in spaces over finite fields. In this chapter, we continue this study, focusing on point–line incidences in finite planes. As we will see, much less is known in finite fields and most incidence problems seem to become more difficult to study. Unlike the case of $\mathbb{R}^d$, we do not have one main technique that leads to most of the current best bounds. Instead, each bound that we derive in this chapter requires a rather different set of tools.

13.1   Preliminaries

As we saw in Chapter 7, the Szemerédi–Trotter theorem can be generalized to $\mathbb{C}^2$. Other works generalize this theorem for semi-algebraic sets, for definable curves and points in o-minimal structures, and more (for example, see [6, 21, 42]). As illustrated by the following example, in $\mathbb{F}_q^2$ the situation is more involved. Let $\mathcal{P}$ be the set of all q^2 points of $\mathbb{F}_q^2$, and let $\mathcal{L}$ be the set of all lines in $\mathbb{F}_q^2$. Each line in $\mathbb{F}_q^2$ is defined by an equation either of the form $y \equiv ax + b$ or of the form $x \equiv b$, where $a, b \in \mathbb{F}_q$. Thus, we have $|\mathcal{L}| = q^2 + q$. Each point of $\mathcal{P}$ is incident to $q + 1$ lines of $\mathcal{L}$ (a line for each possible slope), so

$$I(\mathcal{P}, \mathcal{L}) = \Theta(q^3) = \Theta(|\mathcal{P}|^{3/4}|\mathcal{L}|^{3/4}).$$

The above example shows that there is no hope for obtaining the Szemerédi–Trotter bound $I(\mathcal{P}, \mathcal{L}) = O(|\mathcal{P}|^{2/3}|\mathcal{L}|^{2/3} + |\mathcal{P}| + |\mathcal{L}|)$. We can use a combinatorial argument to get the following weaker bound.

Lemma 13.1. *Let $\mathcal{P}$ be a set of m points and let $\mathcal{L}$ be a set of n lines, both in $\mathbb{F}_q^2$. Then $I(\mathcal{P}, \mathcal{L}) = O(m^{3/4}n^{3/4} + m + n)$.*

Proof. Let $\mathcal{L}_x$ be the set of lines of $\mathcal{L}$ that are defined by an equation of the form $x \equiv b$ for some $b \in \mathbb{F}_q$. That is, $\mathcal{L}_x$ is the subset of lines of $\mathcal{L}$ that are parallel to the x-axis. Since each point of $\mathbb{F}_q^2$ is incident to at most one line of $\mathcal{L}_x$, we have $I(\mathcal{P}, \mathcal{L}_x) \le m$. We may thus remove the lines of $\mathcal{L}_x$ from $\mathcal{L}$, and assume that no line of $\mathcal{L}$ is parallel to the x-axis.

We note that two lines in $\mathbb{F}_q^2$ intersect in at most one point. Indeed, a point on two lines satisfies two equations of the form $y \equiv ax + b$, and such a system has at most one solution. That is, the incidence graph of $\mathcal{P} \times \mathcal{L}$ contains no $K_{2,2}$. Applying Lemma 7.1 gives

$$I(\mathcal{P}, \mathcal{L}) = O(m\sqrt{n} + n). \tag{13.1}$$

We perform the following point-line duality. The *dual of a point* $p = (a,b) \in \mathbb{F}_q^2$ is the line p^* defined by $y \equiv ax - b$. Similarly, the *dual of a line* ℓ defined by $y \equiv cx - d$ is the point $p^* = (c,d) \in \mathbb{F}_q^2$ (recall that we assume that no line of $\mathcal{L}$ is of the form $x \equiv b$). Note that a point p is incident to a line ℓ if and only if the point ℓ^* is incident to the line p^*, since both hold if and only if $d \equiv ac - b$. Consider the point set $\mathcal{L}^* = \{\ell^* : \ell \in \mathcal{L}\}$ and the set of lines $\mathcal{P}^* = \{p^* : p \in \mathcal{P}\}$. By the observation above, we have $I(\mathcal{P}, \mathcal{L}) = I(\mathcal{L}^*, \mathcal{P}^*)$. Thus, by applying Lemma 7.1 on $\mathcal{L}^*$ and $\mathcal{P}^*$ we get

$$I(\mathcal{P}, \mathcal{L}) = I(\mathcal{L}^*, \mathcal{P}^*) = O(n\sqrt{m} + m). \tag{13.2}$$

By multiplying (13.1) and (13.2) we have

$$\begin{aligned} I(\mathcal{P}, \mathcal{L}) \cdot I(\mathcal{P}, \mathcal{L}) &= O\left((m\sqrt{n} + n)(n\sqrt{m} + m)\right) \\ &= O\left(m^{3/2}n^{3/2} + m^2 n^{1/2} + m^{1/2}n^2 + mn\right). \end{aligned} \tag{13.3}$$

If $m = \Omega(n^2)$ then (13.2) implies $I(\mathcal{P}, \mathcal{L}) = m$. We may thus assume that $m = O(n^2)$, which in turn implies $m^2 n^{1/2} = O(m^{3/2}n^{3/2})$. A symmetric argument yields $m^{1/2}n^2 = O(m^{3/2}n^{3/2})$. Also, note that $mn = O(m^2 + n^2)$. Combining these observations with (13.3) gives

$$I(\mathcal{P}, \mathcal{L})^2 = O\left(m^{3/2}n^{3/2} + m^2 + n^2\right).$$

Taking the square root of both sides yields the assertion of the lemma. $\square$

The idea of moving to a dual space is very useful when studying incidences. We already encountered a variant of it in Problem 3.13. We will see two other variants

in Section 13.4 below, another in Problem 13.1, and more examples throughout the following chapters.

The point-line example before Lemma 13.1 shows that the bound of this lemma is tight. This is an extreme example in the sense that it uses all of the points and lines in $\mathbb{F}_q^2$. As the number of points and lines becomes smaller than q^2, better incidence bounds can be obtained. The following result studies the other extreme, where m and n are unusually small compared to q.

Theorem 13.2 (Grosu [47]). *Let p be a prime number. Let $\mathcal{P}$ be a set of n points and let $\mathcal{L}$ be a set of n lines, both in $\mathbb{F}_p^2$. If $n = O(\log\log\log p)$ then $I(\mathcal{P}, \mathcal{L}) = O(n^{4/3})$.*

To see that the bound of Theorem 13.2 is tight, note that we can use Elekes' construction from Claim 1.3. Indeed, in that construction the points have integer coordinates and the lines are defined by equations with integer coordinates. Thus, the same construction exists in $\mathbb{F}_q^2$ as long as q is sufficiently large with respect to m and n.

The above results might have left you confused regarding how a tight point-line incidence bound in $\mathbb{F}_p^2$ should look like. Indeed, for most ranges of m and n this is an open problem. In Sections 13.3 and 13.6 we will study the current best bounds for this problem.

13.2 A brief introduction to the projective plane

Intuitively, the projective plane is an extension of the plane obtained by adding an extra line at infinity. This extension leads to several nice properties, such as that every two lines intersect in a point. While projective planes can be considered with arbitrary fields, we introduce these specifically for the case of finite fields. The projective plane over $\mathbb{R}$ is defined in the same way. When reading the explanations below, one might get additional geometric intuition by also thinking about the case of $\mathbb{R}$. A nice introduction of real projective spaces can be found for example in [23, Chapter 8].

We define the *projective plane* $\mathbb{PF}_q^2$ as follows. The set $\mathbb{PF}_q^2$ consists of the points of $\mathbb{F}_q^3 \setminus \{0\}$, but with two points $u, v \in \mathbb{PF}_q^2$ considered equivalent if there exists $c \in \mathbb{F}_q$ such that $u \equiv cv$. That is, a projective point is an equivalence class of points in $\mathbb{F}_q^3 \setminus \{0\}$. Such an equivalence class is a line incident to the origin in $\mathbb{F}_q^3$ (without the origin itself). Note that the q^2 points of the form $(1, a, b)$ with $a, b \in \mathbb{F}_q$ are all distinct. In addition to these points, $\mathbb{PF}_q^2$ also contains the q distinct points of the form $(0, 1, b)$ with $b \in \mathbb{F}_q$ and the point $(0, 0, 1)$. It can be easily verified that there

are no other points in $\mathbb{PF}_q^2$, so $|\mathbb{PF}_q^2| = q^2 + q + 1$. We refer to the non-projective plane $\mathbb{F}_q^2$ as the *affine plane*.

Denote the coordinates of $\mathbb{PF}_q^2$ as $x, y,$ and z. When working in $\mathbb{PF}_q^2$, we cannot define a variety using an arbitrary set of polynomials. For example, the "line" that is defined by $x + y - 3 \equiv 0$ contains the point $(2, 1, 1)$ but not the point $(4, 2, 2)$. This is impossible, since these are two representations of the same point. To overcome this issue, in $\mathbb{PF}_q^2$ we define varieties only with homogeneous polynomials.[1] For every homogeneous polynomial $f \in \mathbb{F}_q[x, y, z]$ of degree k, point $u \in \mathbb{F}_q^3$, and number $c \in \mathbb{F}_q$, we have $f(c \cdot u) \equiv c^k f(u)$. That is, $f(c \cdot u) \equiv 0$ if and only if $f(u) \equiv 0$, so varieties defined with homogeneous polynomials are consistently defined.

We can move from the affine plane $\mathbb{F}_q^2$ to the corresponding projective plane $\mathbb{PF}_q^2$ by taking a point $(x, y) \in \mathbb{F}_q^2$ to the point $(x, y, 1) \in \mathbb{PF}_q^2$. To move varieties from $\mathbb{F}_q^2$ to $\mathbb{PF}_q^2$, we *homogenize* them as follows. Given a polynomial $f \in \mathbb{F}_q[x, y]$ of degree k, we give every monomial of f degree k by multiplying it with some power of z (multiplying a monomial of degree k' by $z^{k-k'}$). We denote the homogenization of $f \in \mathbb{F}_q[x, y]$ as $f^* \in \mathbb{F}_q[x, y, z]$. Note that f vanishes on a point $(p_x, p_y) \in \mathbb{F}_q^2$ if and only if f^* vanishes on $(p_x, p_y, 1)$, so the transition from $\mathbb{F}_q^2$ to $\mathbb{PF}_q^2$ maintains point-variety incidences.

We now consider a more geometric interpretation of $\mathbb{PF}_q^2$. The process of moving from $\mathbb{F}_q^2$ to $\mathbb{PF}_q^2$ can be thought of as taking the affine plane $\mathbb{F}_q^2$ and placing it in $\mathbb{F}_q^3$ as the plane defined by $z \equiv 0$. Then, a point $u \in \mathbb{PF}_q^2$ becomes the line that is incident to the origin and to u. This process covers all of the points in $\mathbb{F}_q^3$ that have a nonzero z-coordinate. The points with a zero z-coordinate are also lines that pass through the origin, but they do not intersect the plane defined by $z \equiv 1$ and do not correspond to points of $\mathbb{F}_q^2$. These points are on an extra projective line that is defined by $z \equiv 0$. We think of this line as being at infinity, for reasons explained below.

Let ℓ be a line in $\mathbb{F}_q^2$ defined by $ax + by + c \equiv 0$, and let ℓ^* be the corresponding projective line in $\mathbb{PF}_q^2$. That is, ℓ^* is defined by $ax + by + cz \equiv 0$. As explained above, every point of $\mathbb{F}_q^2$ that is incident to ℓ corresponds to a point of $\mathbb{PF}_q^2$ that is incident to ℓ^*. However, ℓ^* also contains the extra point $(1, -b^{-1}a, 0)$ on the projective line at infinity (if $b \equiv 0$, then the extra point is instead $(1, 0, 0)$). This extra point is defined by the slope of ℓ^*, so parallel projective lines intersect in a point at infinity. For example, every projective line with slope 1 can be defined as $y \equiv x + zc$ for some $c \in \mathbb{F}_q$, so all of these lines intersect in $(1, 1, 0)$. In general, any two projective lines in $\mathbb{PF}_q^2$ intersect in exactly one point. One reason for saying that the line $z \equiv 0$ is at infinity, is that in painting and photography parallel lines seem as if they meet at infinity (see Figure 13.1).

---

[1] Recall that a polynomial is homogeneous if all of its monomials are of the same degree.

Figure 13.1: Parallel lines meet at infinity.

In affine planes, we work with transformations that take lines to lines, such as translations, rotations, and scalings. Similarly, *projective transformations* are bijections between $\mathbb{PF}_q^2$ and itself that take lines to lines. We can define such a transformation from (x, y, z) to (x', y', z') as

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{1,2} & a_{2,2} & a_{2,3} \\ a_{1,3} & a_{2,3} & a_{3,3} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{13.4}$$

where $a_{j,k} \in \mathbb{F}_q$ and the 3×3 matrix is invertible.

Recall the geometric interpretation of $\mathbb{PF}_q^2$ as the intersection of the plane $z \equiv 1$ with lines through the origin. From this perspective, a projective transformation corresponds to moving the plane, and thus changing where every line intersects it.

We rely on properties of the projective plane in Sections 13.3 and 13.6. We conclude this section with an observation that will be useful to us below. Consider a point $u \in \mathbb{PF}_q^2$ with a nonzero z-coordinate and let $\mathcal{L}$ be a set of projective lines that are incident to u. Let τ be a projective transformation that takes u to the line at infinity (that is, to a point with a zero z-coordinate). Since τ takes the intersection point of the lines of $\mathcal{L}$ to infinity, all of these lines become parallel.

13.3 Incidences between large sets of points and lines

We now return to studying incidences between m point and n lines in $\mathbb{F}_q^2$. In Section 13.1 we mentioned that when m and n are tiny with respect to q, the standard Szemerédi–Trotter bound $O(m^{2/3}n^{2/3} + m + n)$ holds and is tight (at least when q is a prime and $m \approx n$). We also saw that this bound is false when we take all of the points and all of the lines of $\mathbb{F}_q^2$. There is a large gap between these two cases, and

we currently do not know what the correct bound is for most of it. In this section we derive another incidence bound, which is asymptotically tight when $mn = \Omega(q^3)$. This bound is equivalent to the Szemerédi–Trotter bound when both m and n are $\Theta(q^{3/2})$.

Theorem 13.3. *Let $\mathcal{P}$ be a set of m points and let $\mathcal{L}$ be a set of n lines, both in $\mathbb{F}_q^2$. Then*

$$I(\mathcal{P}, \mathcal{L}) = O\left(\frac{mn}{q} + \sqrt{mnq}\right).$$

When $mn = \Omega(q^3)$, we have the bound $I(\mathcal{P}, \mathcal{L}) = \Theta(mn/q)$ for every such point-line configuration in $\mathbb{F}_q^2$.

Theorem 13.3 has several rather different proofs. There are proofs that rely on the Fourier transform and on basic Additive Combinatorics. Here we present a proof of Vinh [103] that is based on Spectral Graph Theory. Before presenting this proof, we briefly introduce a few concepts from Spectral Graph Theory. A nice introduction to this topic can be found in [105, Section 8.6].

The *adjacency matrix* M of a graph $G = (V, E)$ is a $|V| \times |V|$ matrix that is defined as follows. Let $N = |V|$ and write $V = \{v_1, \ldots, v_N\}$. Then the cell M_{ij} contains the number of edges in E between v_i and v_j. When there are no parallel edges in G, the matrix of M consists only of ones and zeros.

The cell M_{ij}^2 (where $M^2 = M \cdot M$) contains the number of paths of length two in G between v_i and v_j. When there are no loops[2] in G, the number of paths of length two between a vertex to itself is the degree of the vertex, so $M_{ii} = \deg v_i$. To have this property hold also when there are loops in G, we have every loop contribute only one to the degree of the corresponding vertex. That is, the degree of a vertex v is defined as the number of distinct edges that are adjacent to v. If the eigenvalues of M are $\lambda_1, \lambda_2, \ldots, \lambda_N$ then the eigenvalues of M^2 are $\lambda_1^2, \lambda_2^2, \ldots, \lambda_N^2$.

The *eigenvalues of a graph* G are the eigenvalues of the adjacency matrix of G. A graph G is *k-regular* if every vertex of G is of degree k. The largest eigenvalue of a k-regular graph is k. We will rely on the following result concerning eigenvalues of regular graphs (for example, see [2, Corollary 9.2.5]).

Lemma 13.4. *Let $G = (V, E)$ be a k-regular graph with eigenvalues $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{|V|}|$. Let $B, C \subset V$ be (not necessarily disjoint) subsets of vertices of G, and let $e(B, C)$ denote the number of edges in E that have one endpoint in B and the*

---

[2]Recall that a loop is an edge with the same vertex at both of its endpoints.

other endpoint in C. Then

$$\left| e(B,C) - \frac{k}{n}|B||C| \right| \le \lambda_2 \sqrt{|B||C|}.$$

Consider a random graph G that is generated by starting with a vertex set V and inserting each of the $|V|^2$ potential edges (including loops) with probability k/n. When taking two arbitrary vertex subsets $B, C \subset V$, the expected number of edges with one endpoint in B and the other endpoint in C is $\frac{k}{n}|B||C|$. Intuitively, Lemma 13.4 states that a regular graph with only one large eigenvalue is behaves like a random graph in this sense.

Proof of Theorem 13.3. We consider the projective plane $\mathbb{PF}_q^2$, as defined in Section 13.2. Recall that the number of points in $\mathbb{PF}_q^2$ is $q^2 + q + 1$. We construct the graph $G_q = (\mathbb{PF}_q^2, E)$ as follows. The graph contains a vertex for every point of $\mathbb{PF}_q^2$, and $(u, v) \in E$ for vertices $u = (u_1, u_2, u_3)$ and $v = (v_1, v_2, v_3)$ if and only if

$$u_1 v_1 + u_2 v_2 + u_3 v_3 \equiv 0. \tag{13.5}$$

We can think of this condition as asking the point (v_1, v_2, v_3) to be incident to the projective line defined by $u_1 x + u_2 y + u_3 z \equiv 0$. Note that replacing u with an equivalent representation cu (where $c \in \mathbb{F}_q \setminus \{0\}$) does not change the condition (13.5), and similarly for v. That is, (13.5) is well defined with respect to the equivalence relation.

Consider the number of solutions to $u_1 x + u_2 y + u_3 z \equiv 0$ where (u_1, u_2, u_3) and (x, y, z) are in $\mathbb{PF}_q^2$ (and x, y, z are variables). Recall that an equation of the form $ax \equiv b$ has the unique solution $x \equiv a^{-1}b$, unless $a \not\equiv 0$. Without loss of generality, we assume that $u_1 \not\equiv 0$. Then for any choice of $y, z \in \mathbb{F}_q$ there is a unique $x \in \mathbb{F}_q$ that solves $u_1 x + u_2 y + u_3 z \equiv 0$. There are $q^2 - 1$ choices for y and z where not both are zero, and each can be completed to a solution in a unique way. Since there are $q^2 - 1$ solutions and each solution has $q - 1$ equivalent representations, we have $q + 1$ distinct solutions in this case. This implies that G_q is a $(q + 1)$-regular graph (recall that a loop increases the degree of the corresponding vertex by one).

For two distinct points $(u_1, u_2, u_3), (u_1', u_2', u_3') \in \mathbb{PF}_q^2$, the system of equations $u_1 x + u_2 y + u_3 z \equiv 0$ and $u_1' x + u_2' y + u_3' z \equiv 0$ has a unique solution in $\mathbb{PF}_q^2$ (that is, it has $q - 1$ equivalent solutions). That is, any two vertices of G_q have exactly one path of length two between them.

Let M be the adjacency matrix of G_q, and consider M^2. Recall that M_{ij}^2 is the number of paths of length two between the vertices v_i and v_j. Thus, by the preceding

paragraph every cell of M^2 that is not on the main diagonal contains 1. Since G_q is $(q + 1)$-regular, the values on the main diagonal of M are all $q + 1$. That is,

$$M_{ij}^2 = \begin{cases} 1, & i \neq j, \\ q + 1, & i = j. \end{cases} \tag{13.6}$$

An all 1's matrix of size $N \times N$ has the eigenvalue N with multiplicity one and the eigenvalue 0 with multiplicity $N - 1$. Indeed, the eigenvalue N has the eigenvector $(1, 1, \ldots, 1)$, and the eigenvalue 0 has the eigenvectors $(1, -1, 0, \ldots, 0)$, $(1, 0, -1, 0, \ldots, 0)$, and so on. Increasing every element on the main diagonal of a matrix by c shifts all of the eigenvalues of that matrix by c. Combining these two properties with (13.6) implies that M^2 has the eigenvalue $(q^2 + q + 1) + q = (q + 1)^2$ with multiplicity 1 and the eigenvalue q with multiplicity $q^2 + q$. This in turn implies that the absolute value of one eigenvalue of M is $q + 1$ and the absolute values of the other eigenvalues are $\sqrt{q}$.

We consider every point $a = (a_x, a_y) \in \mathcal{P}$ as the point $(a_x, a_y, 1) \in \mathbb{PF}_q^2$, and note that no two points of $\mathcal{P}$ are equivalent after this change. We consider a line $\ell \in \mathcal{L}$ defined by $b_y y + b_x x + b_1 \equiv 0$ as the point $(b_x, b_y, b_1) \in \mathbb{PF}_q^2$. Once again, two distinct lines of $\mathcal{L}$ cannot go to equivalent points of $\mathbb{PF}_q^2$. Note that the point a is incident to the line ℓ if and only if $a_x b_x + a_y b_y + b_1 \equiv 0$, which is equivalent to G_q containing an edge between the vertex of a and the vertex of ℓ. Thus, $I(\mathcal{P}, \mathcal{L}) = e(\mathcal{P}, \mathcal{L})$. We apply Lemma 13.4 with $B = \mathcal{P}$, $C = \mathcal{L}$, and $|\lambda_2| = \sqrt{q}$, to obtain

$$\left| I(\mathcal{P}, \mathcal{L}) - \frac{q + 1}{q^2 + q + 1} mn \right| = \left| e(\mathcal{P}, \mathcal{L}) - \frac{q + 1}{q^2 + q + 1} mn \right| \leq \sqrt{qmn}.$$

The above equation immediately implies the asserted bound $I(\mathcal{P}, \mathcal{L}) = O(mn/q + \sqrt{qmn})$. When $mn = \Omega(q^3)$ we have that $\frac{q+1}{q^2+q+1} mn = \Omega(\sqrt{qmn})$, and thus $I(\mathcal{P}, \mathcal{L}) = \Theta(mn/q)$. That is, not only we get that the incidence bound is tight in this range — we get that *every configuration* of m points and n lines has $\Theta(mn/q)$ incidences. $\square$

Theorem 13.3 provides a tight bound when $mn = \Omega(q^3)$, and implies the Szemerédi–Trotter bound when both m and n are $\Theta(q^{3/2})$. However, the case of smaller m and n remains wide open. In Section 13.6, we will study the current best bounds in this range.

13.4 Incidences with planes in $\mathbb{F}_q^3$

In this section we study a bound of Rudnev [81] for point–plane incidences in $\mathbb{F}_q^3$. In some sense, Rudnev's bound is currently our strongest tool for handling problems with sets that are not very large. The current best bounds for many finite fields problems, such as the sum-product problem and point–line incidences in the plane, are all obtained via reductions to Rudnev's bound.

When considering incidences in $\mathbb{F}_q^d$ where $d \geq 3$, we have the same issue as in $\mathbb{R}^d$: By placing the points on a line ℓ and taking planes that contain ℓ, we get that every plane is incident to every point. That is, the problem is trivial. In this section, we avoid the above issue by assuming that the no line in $\mathbb{F}_q^3$ contains k points. We begin by deriving a weak bound for the problem, relying on a standard combinatorial argument.

Lemma 13.5. *Let $\mathcal{P}$ be a set of m points and let Π be a set of n planes, both in $\mathbb{F}_q^3$. If no line in $\mathbb{F}_q^3$ contains k points of $\mathcal{P}$, then*

$$I(\mathcal{P}, \Pi) = O(n\sqrt{km} + m).$$

Proof sketch. By the assumption involving k, the incidence graph of $\mathcal{P} \times \Pi$ contains no $K_{k,2}$. By noting that the proof of Lemma 3.4 remains valid also for point-plane incidences in $\mathbb{F}_q^3$ with no $K_{s,t}$ in the incidence graph. By revising this proof to include the exact dependency in s and t, we obtain the bound $I(\mathcal{P}, \Pi) = O(mt^{1/s}n^{1-1/s} + sn)$.

We perform a point-plane duality, imitating the point-line duality from the proof of Lemma 13.1. This leads to a set of n points and a set of m lines in $\mathbb{F}_q^3$, with no $K_{t,s}$ in the incidence graph. Applying the above bound after the duality argument gives $I(\mathcal{P}, \Pi) = O(ns^{1/t}m^{1-1/t} + tm)$. To complete the proof we set $s = k$ and $t = 2$. $\square$

Rudnev [81] derived the following bound for the point–plane incidence problem.

Theorem 13.6. *Let $q = p^r$ for some prime p and positive integer r. Let $\mathcal{P}$ be a set of m points and let Π be a set of n planes, both in $\mathbb{F}_q^3$, such that $n \geq m$. Assume that $m = O(p^2)$ and that no line of $\mathbb{F}_q^3$ contains k points of $\mathcal{P}$. Then*

$$I(\mathcal{P}, \Pi) = O(n\sqrt{m} + kn).$$

At first, Theorem 13.6 might not look very impressive. This theorem improves upon Lemma 13.5 only by having a better dependency on k, while also adding two new restrictions. However, this improved dependency on k turned out to be the main

tool for deriving the current best bounds for many problems. We will see impressive uses of Rudnev's theorem in Sections 13.5 and 13.6.

We present a simplified proof of Theorem 13.6, by de Zeeuw [110]. When working in $\mathbb{F}_q^d$, it is sometimes easier to prove an argument in the algebraic closure of $\mathbb{F}_q$. In our proof of Theorem 13.6, we are going to move to the algebraic closure. We do not go over the full details of how this step works, to avoid many technicalities. The algebraic closure will be used in the following proof when studying properties of ruled surfaces, when performing a generic rotation of the space, and when bounding the number of lines in the intersection of two surfaces.

In $\mathbb{R}^3$ or $\mathbb{C}^3$, we can derive point–plane incidence bounds by using polynomial partitioning. Unfortunately, this tool is not available to us when working in $\mathbb{F}_q^3$. Rudnev cleverly avoided this issue by reducing the problem to bounding the number of intersecting pairs in a set of lines in $\mathbb{F}_q^3$. Recall that the distinct distances problem was reduced to a line-intersection problem (see Chapter 8). In Chapter 9 we solved this line intersection problem by using polynomial partitioning. However, the original proof of Guth and Katz [51] was based on different ideas, which extend more easily to $\mathbb{F}_q^3$.

To describe Rudnev's reduction, we first need some notation. We use the coordinates $(x, y, z) \in \mathbb{F}_q^3$, and denote the x-coordinate of a point $p \in \mathbb{F}_q^3$ as p_x. We require the following objects for the reduction.

- H_1: the plane defined by $x \equiv 1$.
- ℓ_z: the z-axis of $\mathbb{F}_q^3$. That is, the line defined by $x \equiv 0$ and $y \equiv 0$.
- $\mathcal{L}$: the set of lines in $\mathbb{F}_q^3$ that intersect both ℓ_z and H_1.
- D: The set of points and planes in $\mathbb{F}_q^3$ defined as follows. A point $p \in \mathbb{F}_q^3$ is in D if it satisfies $p_x \not\equiv 0$. A plane $H \subset \mathbb{F}_q^3$ is in D if H intersects both H_1 and ℓ_z but does not contain ℓ_z

For example, note that $\ell_z \notin \mathcal{L}$ since it does not intersect H_1. Similarly, note that no line of $\mathcal{L}$ is contained in H_1.

Lemma 13.7. *There exists a map ϕ that takes every element of $\mathcal{P} \cup \Pi$ to a line in $\mathbb{F}_q^3$, such that $p \in \mathcal{P}$ is incident to $H \in \Pi$ if and only if the lines $\phi(p)$ and $\phi(H)$ intersect.*

Proof. For every line $\ell \in \mathcal{L}$, there exist $a, b, c \in \mathbb{F}_q$ such that $\ell \cap \ell_z = (0, 0, a) \in \mathbb{F}_z^3$ and $\ell \cap H_1 = (1, b, c) \in \mathbb{F}_q^3$. We define the *dual* of ℓ as the point $\ell^* = (a, b, c) \in \mathbb{F}_z^3$. For a plane $H \in \Pi$, we set

$$\phi(H) = \{\ell^* : \ell \in \mathcal{L} \text{ and } \ell \subset H\}.$$

If H intersects ℓ_z at $(0, 0, z_H)$, then the x-coordinate of every element of $\phi(H)$ is z_H. Since $H \cap H_1$ is a line, we can write $H \cap H_1 = \{(1, t, u_H t + v_H) : t \in \mathbb{F}_q\}$ for

some $u_H, v_H \in \mathbb{F}_q^3$. Every line of $\mathcal{L}$ in H intersects H_1 in a point of this line. This implies that $\phi(H)$ is the line $\{(z_H, t, u_H t + v_H) : t \in \mathbb{F}_q\}$.

For $p \in \mathcal{P}$ we set

$$\phi(p) = \{\ell^* : \ell \in \mathcal{L} \text{ and } p \in \ell\}.$$

All of the lines that are incident to p and intersect ℓ_z have the same projection on the xy-plane. This implies that all of these lines intersect H_1 at points with the same y-coordinate. Thus, the y-coordinates of all the points of $\phi(p)$ are the same. Denote this y-coordinate as y_p.

Let ℓ be a line that is incident to p and intersects ℓ_z at $(0, 0, t)$. The projection of ℓ on the xz-plane is linearly determined by t. For example, the slope of the projected line is $(p_z - t)/(p_x - 0)$. Thus, there exist $u_p, v_p \in \mathbb{F}_q^3$ such that ℓ intersect H_1 at $(1, y_p, u_p t + v_p)$. We conclude that $\phi(p)$ is the line $\{(t, y_p, u_p t + v_p) : t \in \mathbb{F}_q\}$.

Now that we have a well defined map $\phi(\cdot)$, it remains to prove the incidence property of the lemma. Consider a point $p \in \mathcal{P}$ and a plane $H \in \Pi$. Note that $p \in H$ if and only if there exists a line $\ell \in \mathcal{L}$ that is incident to p and contained in H. A line $\ell \in \mathcal{L}$ satisfies the above properties if and only if $\ell^* \in \phi(p) \cap \phi(H)$. That is, $p \in H$ if and only if $\phi(p) \cap \phi(H) \neq \emptyset$. $\qquad\square$

The above reduction is somewhat surprising, since it is unclear where it came from. Rudnev's original reduction took each line to a point on the *Klein quadric* and studied properties of this object. De Zeeuw simplified the proof, partly by removing any statements involving the Klein quadric. This led to the unusual formulation above.

Ruled surfaces. Up to a few technicalities, Lemma 13.7 provides a reduction from the point-plane incidence problem to a line intersection problem. To solve this line intersection problem, we require some properties of ruled surfaces.

As with other fields, for a polynomial $f \in \mathbb{F}_q[x_1, x_2, x_3]$, we define

$$\mathbf{V}(f) = \{(a_1, a_2, a_3) \in \mathbb{F}_q^3 : f(a_1, a_2, a_3) \equiv 0\}.$$

Since we are working in a three-dimensional space and with a single polynomial, we refer to $\mathbf{V}(f)$ as a *surface*. The *degree* of a surface S is the minimum degree of a polynomial $f \in \mathbb{F}_q[x_1, x_2, x_3]$ such that $\mathbf{V}(f) = S$.

An irreducible surface S in $\mathbb{F}_q^3$ (or in any other three-dimensional space, such as $\mathbb{R}^3$) is *ruled* if for every point $p \in S$ there exists a line that is contained in S and incident to p. Some examples of ruled surfaces are planes, cylinders, and conical surfaces. A surface S is said to be *doubly-ruled* if for every point $p \in S$ there exist two lines that

are contained in S and incident to p. Similarly, a plane is said to be *infinitely-ruled*, and ruled surfaces that are not doubly-ruled are said to be *singly-ruled*.

We now state some properties of ruled surfaces in $\mathbb{F}_q^3$ without proofs. In Chapter ??? we will discuss ruled surfaces in $\mathbb{R}^3$ and provide proofs for some of these properties. For the case of finite fields, see for example [59, 81]. Recall that this is one of the places where we rely on working in the algebraic closure of $\mathbb{F}_q$.

Lemma 13.8. *For $q = p^r$, let $S \subset \mathbb{F}_q^3$ be an irreducible surface of degree D and let $p \neq 2$.*
(a) If for every point $p \in S$ there exist at least three lines that are contained in S and incident to p, then S is a plane.
(b) If S is doubly-ruled but not a plane, then the lines that are contained in S can be partitioned into two sets R_1, R_2 that satisfy the following properties. The lines of each R_j are a ruling of S. That is $\cup_{\ell \in R_j} \ell = S$, for $j \in \{1, 2\}$. Two lines that are contained in S intersect if and only if they are in different rulings of S.
(c) If S is singly-ruled then it contains at most two special lines that intersect an infinite number of other lines contained in S. Every non-special line in S intersects at most D non-special lines contained in S.
(d) If S is not ruled then it contains $O(D^2)$ lines.

We will require the following variant of Lemma 6.3.

Lemma 13.9. *Let $\mathcal{L}$ be a set of n lines in $\mathbb{F}_q^3$, where $n \leq (q-1)^2/6$. Then there exists a surface $S \subsetneq \mathbb{F}_q^3$ of degree at most $\sqrt{6n} \leq q - 1$ that contains every line of $\mathcal{L}$.*

Proof. Our approach is similar to the one we used in the proof of Lemma 5.4. Consider a polynomial $f \in \mathbb{F}_q[x_1, x_2, x_3]$ of degree at most $\sqrt{6n}$, and denote the coefficients of the monomials of f as $c_1, \ldots, c_k$. The number of distinct monomials in $\mathbb{F}_q[x_1, x_2, x_3]$ of degree at most $\sqrt{6n}$ is $\binom{\sqrt{6n}+3}{3}$, so $k = \binom{\sqrt{6n}+3}{3}$. Write $f = \sum_{j=0}^{\sqrt{6n}} f_j$, where f_j is a homogeneous polynomial of degree j.

We can write a line $\ell \in \mathcal{L}$ as $\{(a_1 t + b_1, a_2 t + b_2, a_3 t + b_3) : t \in \mathbb{F}_q\}$, for some parameters $a_1, a_2, a_3, b_1, b_2, b_3 \in \mathbb{F}_q$. Since the degree of f is at most $\sqrt{6n} \leq q - 1$, by the Schwartz–Zippel lemma (Lemma 6.2) we have that $\ell \subset \mathbf{V}(f)$ if and only if $f(a_1 t + b_1, a_2 t + b_2, a_3 t + b_3) \equiv 0$. That is, if and only if for every $0 \leq j \leq \sqrt{6n}$, the coefficient of t^j in $f(a_1 t + b_1, a_2 t + b_2, a_3 t + b_3)$ is 0. This is a system of $\sqrt{6n} + 1$ linear homogeneous equations in the coefficients c_j.

Asking $\mathbf{V}(f)$ to contain all n lines of $\mathcal{L}$ yields a system of $n(\sqrt{6n} + 1)$ linear homogeneous equations in $\binom{\sqrt{6n}+3}{3}$ variables. Since $\binom{\sqrt{6n}+3}{3} > n(\sqrt{6n}+1)$, we get that this system has a nontrivial solution (it is not difficult to verify that this argument

still holds when working over finite fields). That is, there exists a nonzero polynomial f of degree at most $\sqrt{6n}$ such that $\mathbf{V}(f)$ contains all of the lines of $\mathcal{L}$. By the Schwartz–Zippel lemma, f does not vanish on all of $\mathbb{F}_q^3$. $\qquad\square$

We will also require the following result by Kollár [59, Corollary 40].

Lemma 13.10. *Let $q = p^r$ and let $\mathcal{L}$ be a set of n lines in $\mathbb{F}_q^3$, such that $n = O(p^2)$. If every surface of degree one or two contains $O(\sqrt{n})$ lines of $\mathcal{L}$, then the number of points in $\mathbb{F}_q^3$ that are incident to at least two lines of $\mathcal{L}$ is $O(n^{3/2})$.*

We are finally ready to prove the main result of this section.

Proof of Theorem 13.6. If $p = 2$ then by an assumption of the theorem the number of points is bounded by some constant. This case is easily handled by taking the constant in the $O(\cdot)$-notation of the incidence bound to be sufficiently large. We may thus assume that $p \neq 2$, which allows us to rely on Lemma 13.8. The rest of the proof consists of two parts: reducing the problem to a line intersection problem, and then solving the line intersection problem.

Reducing the problem. We would like to use the reduction from Lemma 13.7, with $\mathcal{L} = \{\phi(p) : p \in \mathcal{P}\}$ and $\mathcal{L}' = \{\phi(H) : H \in \Pi\}$. Recall that this lemma only holds for points with a nonzero x-coordinate and for planes that intersect H_1 and ℓ_z. This issue can be easily resolved by permuting the three coordinates x, y, z and then performing a translation in the new x-direction. However, we would also like to argue that no two lines of $\mathcal{L}$ intersect, and that no two lines of $\mathcal{L}'$ intersect. To obtain all of the above properties simultaneously, we perform a generic rotation of $\mathbb{F}_q^3$. More precisely, we perform this generic rotation in the three-dimensional space over the algebraic closure of $\mathbb{F}_q$.

After performing the generic rotation, we can apply the reduction from Lemma 13.7 with $\mathcal{P}$ and Π. Note that $|\mathcal{L}| = m$ and $|\mathcal{L}'| = n$. By Lemma 13.7, the number of incidences $I(\mathcal{P}, \Pi)$ is equal to the number of pairs of intersecting lines in $\mathcal{L} \times \mathcal{L}'$.

Bounding the number of intersecting lines. To complete the proof, it remains to show that the number of intersecting pairs of lines in $\mathcal{L} \times \mathcal{L}'$ is $O(n\sqrt{m} + kn)$. We would like to apply Lemma 13.9, to obtain a surface $S \subsetneq \mathbb{F}_q^3$ of degree $O(\sqrt{m})$ that contains all of the lines of $\mathcal{L}$. This is immediate if $m \leq (q-1)^2/6$. However, we only have the slightly weaker restriction $m \leq cq^2$ for some constant c. We keep the $(q-1)^2/6$ lines of $\mathcal{L}$ that maximize the number of intersections with lines of $\mathcal{L}'$, and discard the other lines of $\mathcal{L}$. This decreases the number of pairs of intersecting lines by a factor of $O_c(1)$, and allows us to apply Lemma 13.9.

We decompose S into irreducible components $S_1, \ldots, S_k$ (that is, we factor a minimum degree polynomial that defines f into irreducible factors and consider the zero set of each). We first consider intersecting pairs of lines from $\mathcal{L} \times \mathcal{L}'$ such that there exists a component S_j that contains exactly one of these two lines. We refer to such an intersection as a *type (i) intersection*. Consider a line $\ell \in \mathcal{L}$ and let S_ℓ be the surface S after removing every component that contains ℓ. Since $\deg S_\ell \leq \deg S = O(\sqrt{m})$, by applying the Schwartz–Zippel lemma (Lemma 6.2) inside of ℓ we obtain $|\ell \cap S_\ell| = O(\sqrt{m})$. Since the lines of $\mathcal{L}'$ are pairwise disjoint, each intersection of ℓ with S_ℓ corresponds to at most one type (i) intersection. A symmetric argument applies when ℓ is a line of $\mathcal{L}'$. By summing the above over every line of $\mathcal{L} \cup \mathcal{L}'$, we conclude that the number of type (i) intersections is $O((m+n)\sqrt{m}) = O(n\sqrt{m})$.

It remains consider intersecting pairs of lines from $\mathcal{L} \times \mathcal{L}'$ such that the two lines are contained in the same set of components of S. We refer to such an intersection as a *type (ii) intersection*. For every line $\ell \in \mathcal{L} \cup \mathcal{L}'$, we *assign* ℓ to the component S_j with the smallest index j among the irreducible components of S that contain ℓ. Let $\mathcal{L}_j$ be the set of lines of $\mathcal{L}$ that are assigned to S_j, and let $\mathcal{L}'_j$ be the set of lines of $\mathcal{L}'$ that are assigned to S_j. Note that if a pair of lines form a type (ii) intersection then both lines are assigned to the same component. Thus, it suffices to bound the number of intersecting pairs of lines in $\mathcal{L}_j \times \mathcal{L}'_j$ for every $1 \leq j \leq k$. Note that $\sum_{j=1}^{k} |\mathcal{L}_j| = m$ and $\sum_{j=1}^{k} |\mathcal{L}'_j| \leq n$. Let I_j denote the number pairs of intersecting lines in $\mathcal{L} \times \mathcal{L}'$ such that both lines are assigned to S_j.

Consider the case where S_j is a plane. Since the lines of $\mathcal{L}_j$ are pairwise disjoint they must all be parallel, and similarly for the lines of $\mathcal{L}'_j$. If the lines of $\mathcal{L}_j$ and $\mathcal{L}'_j$ are all parallel then $I_j = 0$, so we assume that they are not. In this case, every line of $\mathcal{L}_j$ intersects every line of $\mathcal{L}'_j$. By the assumption that no line of $\mathbb{F}_q^3$ contains k points of $\mathcal{P}$, either $|\mathcal{L}_j| < k$ or $|\mathcal{L}'_j| < 2$. Thus, we have $I_j \leq |\mathcal{L}_j| + (k-1)|\mathcal{L}'_j|$. By summing this bound over every irreducible component of S that is a plane, we get $O(m + kn) = O(kn)$ intersecting pairs.

Next, consider the case where S_j is doubly-ruled, and denote the two rulings of S_j as R_1 and R_2. Recall from Lemma 13.8 that two lines of S_j intersect if and only if they are in different rulings of S_j. Since the lines of $\mathcal{L}$ do not intersect, either R_1 or R_2 contain no lines of $\mathcal{L}_j$. Without loss of generality, assume that R_2 contains no lines of $\mathcal{L}_j$. Since no line of $\mathbb{F}_q^3$ contains k points of $\mathcal{P}$, either R_1 contains at most k lines of $\mathcal{L}_j$ or R_2 contains at most one line of $\mathcal{L}'_j$. Thus, we again have $I_j \leq |\mathcal{L}_j| + (k-1)|\mathcal{L}'_j|$. By summing this bound over every irreducible component of S that is doubly-ruled, we get $O(m + kn) = O(kn)$ intersecting pairs.

We move to consider the case where S_j is singly-ruled. By Lemma 13.8, S_j contains

at most two special lines that intersect an infinite number of other lines contained in S_j. Every non-special line in S_j intersects at most $\deg S_j$ non-special lines contained in S_j. The two special lines participate in $O(|\mathcal{L}_j| + |\mathcal{L}'_j|)$ intersections with other lines in S_j. Any other line of $\mathcal{L}_j$ intersects at most $\deg S_j$ non-special lines of $\mathcal{L}'_j$, with a total of $O(\deg S_j |\mathcal{L}_j|)$ such intersections. That is, $I_j = O(|\mathcal{L}'_j| + |\mathcal{L}_j| \deg S_j)$. Since $\sum_j \deg S_j = O(\sqrt{m})$, summing the above bound over every irreducible component of S that is singly-ruled, we get $O(m + n + m^{3/2}) = O(n\sqrt{m})$ intersecting pairs.

It remains to consider the case where S_j is not ruled. Let $\mathcal{L}_j^* = \mathcal{L}_j \cup \mathcal{L}'_j$. By Lemma 13.8, we have $|\mathcal{L}_j^*| = O(\deg S_j^2)$. We now verify that the conditions of Lemma 13.10 hold for $\mathcal{L}_j^*$. We add generic lines to $\mathcal{L}_j^*$, to obtain $|\mathcal{L}_j^*| = \Theta(d_j^2)$. Since $\deg S_j \leq \deg S = O(\sqrt{m})$ and by the assumption that $m = O(p^2)$, we indeed have $|\mathcal{L}_j^*| = O(p^2)$. As in Problem 4.2, if two surfaces of degrees D and E in $\mathbb{F}_q^3$ have no common factors, their intersection contains at most DE lines.[3] Thus, any surface of degree one or two contains $O(\deg S_j)$ lines of $\mathcal{L}_j \cup \mathcal{L}'_j$. By definition no such surface contains more than a few of the additional generic lines. We may thus apply Lemma 13.10, and obtain that $O(\deg S_j^3)$ points of $\mathbb{F}_q^3$ are incident to more than one line of $\mathcal{L}_j^*$. Since the lines of $\mathcal{L}_j$ and disjoint and also the lines of $\mathcal{L}'_j$, every such point corresponds to at most one intersecting pair of $\mathcal{L}_j \cup \mathcal{L}'_j$. That is, $I_j = O(\deg S_j^3)$. By summing this bound over every irreducible component of S that is not ruled, we get

$$\sum_j O(\deg S_j^3) = O\left(\left(\sum_j \deg S_j\right)^3\right) = O(m^{3/2}) = O(n\sqrt{m})$$

intersecting pairs.

There is one small issue that we ignored in the previous paragraph — if S_j is of degree two then might violate the condition of Lemma 13.10 concerning surfaces of degree two. To avoid this issue, note that in this case $|\mathcal{L}_j \cup \mathcal{L}'_j| = O(1)$ so we still have $O(d_j^3) = O(1)$ pairs of intersecting lines.

By Lemma 13.8, the above cases cover all of the irreducible components of S. In conclusion, by summing all of the above cases we have $O(n\sqrt{m} + nk)$ pairs of intersecting lines in $\mathcal{L} \times \mathcal{L}'$. This completes the proof of the theorem. $\qquad \square$

---

[3]To formally prove this we need to rely on the fact that a variant of Bézout's theorem still holds in the algebraic closure of $\mathbb{F}_q^2$.

13.5 The sum-product problem in finite fields

As our first application of Theorem 13.6, we study the sum-product problem in $\mathbb{F}_q$. This problem was already discussed over $\mathbb{R}$ in Section 1.8. Briefly, the problem conjectures that every finite set A satisfies that $\max\{|A + A|, |AA|\}$ is large with respect to $|A|$. Similarly to the case of incidences in finite fields, surprising results are obtained when $A \subset \mathbb{F}_q$ is large with respect to q. In particular, when $A = \mathbb{F}_q$ we have $|A| = |A + A| = |AA| = q$. The problem becomes more interesting when $|A|$ is much smaller than q. The following recent bound is by Roche-Newton, Rudnev, and Shkredov [78].

Theorem 13.11. *Let $q = p^r$ and let $A \subset \mathbb{F}_q$ satisfy $|A| \leq p^{5/8}$. Then*

$$\max\{|A + A|, |AA|\} = \Omega\left(n^{6/5}\right).$$

The current best bound of $\Omega(n^{39/32})$ was obtained in [20] specifically for prime fields. As with the point-line incidences, these bounds are significantly weaker than current bound $\Omega(n^{4/3+\varepsilon})$ over $\mathbb{R}$.

Most of the applications of Theorem 13.6 reduce a problem to point–plane incidences by using the concept of energy. We thus begin by studying this object (already briefly mentioned in Section 12.2).

Given a finite set $A \subset \mathbb{R}$, the *additive energy* of A is

$$E(A) = \left|\left\{(a_1, a_2, a_3, a_4) \in A^4 : a_1 + a_2 = a_3 + a_4\right\}\right|.$$

Note that this is somewhat similar to the set of quadruples from the distinct distances problem. One might refer to that set of quadruples as the distance energy of a point set.

Note that $|A|^2 \leq E(A)$, since this is the number of solutions with $a_1 = a_3$ and $a_2 = a_4$. Similarly, $E(A) \leq |A|^3$ holds since for any choice of a_1, a_2, and a_3 there is at most one valid choice for a_4.

For $x \in A + A$, set $r_A(x) = |\{(a_1, a_2) \in A^2 : a_1 + a_2 = x\}|$. Since every pair of A^2 contributes to exactly one $r_A(x)$, we have that $\sum_{x \in A+A} r_A(x) = |A|^2$. By the Cauchy-Schwarz inequality, we have

$$E(A) = \sum_{x \in A+A} r_A(x)^2 \geq \frac{\left(\sum_x r_A(x)\right)^2}{|A + A|} = \frac{|A|^4}{|A + A|}. \tag{13.7}$$

At first it might seem that the smaller $|A + A|$ is, the larger $E(A)$ is. A few examples to illustrate this:

- If A is a arithmetic progression then $|A + A| = 2|A| - 1$ and $E(A) = \Theta(|A|^3)$.
- If A is a random set, then we expect $|A + A| = \Theta(|A|^2)$ and $E(A) = \Theta(|A|^2)$.
- Let $0 < \alpha < 1$. Set $A = H + R$ with H being an arithmetic progression of size $\Theta(|A|^\alpha)$ and R being a random set of size $\Theta(|A|^{1-\alpha})$. Then $|A + A| = \Theta(|A|^{2-\alpha})$ and $E(A) = \Theta(|A|^{2+\alpha})$.

By (13.7), a small value of $|A + A|$ always implies a large value of $E(A)$. However, the other direction does not always hold. For example, take $A = P \cup R$ where P is a arithmetic progression of size $|A|/2$ and R is a random set of size $|A|/2$. In this case we have $|A + A| = \Theta(|A|^2)$ and $E(A) = \Theta(|A|^3)$.

Before proving our sum-product bound, we first repeat the statement of this result.

Theorem 13.11. *Let q be an odd prime power and let $A \subset \mathbb{F}_q$ satisfy $|A| \le p^{5/8}$. Then*
$$\max\{|A + A|, |AA|\} = \Omega\left(|A|^{6/5}\right).$$

Proof. If $0 \in A$, we remove zero from A. This does not change the asymptotic size of A. Set $A^{-1} = \{1/a : a \in A\}$. We rewrite
$$E(A) = |A|^{-2} \left|\{a_1 + a_2 a_3 / a_3 = a_4 + a_5 a_6 / a_6 \ : \ a_1, \ldots, a_6 \in A\}\right|$$
$$\le |A|^{-2} \left|\left\{(a_1, b_1, c_1, a_2, b_2, c_2) \in \left(A \times AA \times A^{-1}\right)^2 : a_1 + b_1 c_1 = a_2 + b_2 c_2\right\}\right|.$$
$$\tag{13.8}$$

With the above in mind, we define the energy variant
$$E'(A) = \left|\left\{(a_1, b_1, c_1, a_2, b_2, c_2) \in \left(A \times AA \times A^{-1}\right)^2 : a_1 + b_1 c_1 = a_2 + b_2 c_2\right\}\right|.$$

We study $E'(A)$ with a three-dimensional variant of Elekes's sum-product argument (see Section 1.8). Consider the point set
$$\mathcal{P} = \{(a_1, b_2, c_1) \in A \times AA \times A^{-1}\},$$

and the set of planes
$$\Pi = \{x + b_1 z - c_2 y = a_2 \ : \ (a_2, b_1, c_2) \in A \times AA \times A^{-1}\}.$$

A 6-tuple $(a_1, b_1, c_1, a_2, b_2, c_2)$ contributes to $E'(A)$ if and only if the point (a_1, b_2, c_1) is incident to the plane defined by (a_2, b_1, c_2). That is, $E'(A) = I(\mathcal{P}, \Pi)$. We wish to

bound the number of incidences using Theorem 13.6, and first check the conditions of this theorem hold.

We have $m = |\mathcal{P}| = |\Pi| = |A|^2|AA|$. The maximum number of collinear points in $\mathcal{P}$ is $|AA|$, since this is the size of the longest side of the cartesian product $A \times AA \times A^{-1}$. We may assume that $|AA| = O(|A|^{6/5})$, since otherwise we are done. Combining this with $|A| = O(p^{5/8})$ implies that $m = |A|^2|AA| = O(p^2)$. We can thus apply Theorem 13.6 on $\mathcal{P}$ and Π with $k = |AA|$, obtaining

$$E'(A) = I(\mathcal{P}, \Pi) = O\left(m^{3/2} + mk\right) = O\left(|A|^3|AA|^{3/2} + |A|^2|AA|^2\right)$$
$$= O\left(|A|^3|AA|^{3/2}\right).$$

Combining the above with (13.8) gives

$$E(A) \le |A|^{-2}E'(A) = O\left(|A||AA|^{3/2}\right) = O\left(|A||AA|^{3/2}\right) = O\left(|A|^{14/5}\right).$$

Recalling (13.7) leads to $|A + A| = \Omega(|A|^{6/5})$, which completes the proof. $\qquad\square$

With some more work, one can improve the value of k in the proof of Theorem 13.11. Surprisingly, a better bound on k does not improve the theorem in any way.

13.6 Incidences between medium-sized sets of points and lines

In this section we study the current best bound for point-line incidences in $\mathbb{F}_q^2$ for the case where the number of points and lines is not very large and not extremely small. The following result is by Stevens and de Zeeuw [93].

Theorem 13.12. *For $q = p^r$, let $\mathcal{P}$ be a set of m points and $\mathcal{L}$ be a set of n lines, both in $\mathbb{F}_q^2$, such that $m^{7/8} < n < m^{8/7}$ and $n^{13}/m^2 = O(p^{15})$. Then*

$$I(\mathcal{P}, \mathcal{L}) = O(m^{11/15}n^{11/15}).$$

This is another example for how little we currently know about incidences in spaces over finite fields. Note that the bound $O(m^{11/15}n^{11/15})$ is relatively close to the elementary upper bound from Lemma 13.1. Specifically, $3/4 - 11/15 = 1/60$ while $3/4 - 2/3 = 1/12$. Moreover, when $m = o(n^{7/8})$ or $n = o(m^{7/8})$ the elementary bounds (13.1) and (13.2) are stronger. This seems to be part of a general phenomena,

where such combinatorial problems are significantly more difficult when working over a finite field.

The proof of Theorem 13.12 is based on the observation that it is easier to bound the number of incidences when the point set is a lattice. Indeed, in this case we can use an energy argument similar to the one in the proof of Theorem 13.11.

Theorem 13.13. *For $q = p^r$, let $\mathcal{L}$ be a set of n lines in $\mathbb{F}_q^2$. Let $A, B \subset \mathbb{F}_q$ such that $a = |A|$, $b = |B|$, $a \leq b$, $ab^2 = O(n^3)$ and $an = O(p^2)$. Then*

$$I(A \times B, \mathcal{L}) = O(a^{3/4}b^{1/2}n^{3/4} + n).$$

To put this result in context, note that when $a = b = \sqrt{m}$ Theorem 13.13 gives the bound $O(\mathcal{P}, \mathcal{L}) = O(m^{5/8}n^{3/4} + n)$. This improves the bound of Lemma 13.1 by a factor of $m^{1/8}$ (in the special case where the point set is a lattice), which is a significant improvement for an incidence bound.

Proof of Theorem 13.13. We begin with two steps of pruning the set of lines $\mathcal{L}$. In the first step, we remove all of the vertical and horizontal lines from $\mathcal{L}$ (lines defined by $x \equiv c$ for some $c \in \mathbb{F}_q$). Since every point of $A \times B$ is incident to at most one vertical line and to at most one horizontal line, this decreases the number of incidences by at most ab. The assumption $ab^2 = O(n^3)$ is equivalent to $a^{1/4}b^{1/2} = O(n^{3/4})$, which implies $ab = O(a^{3/4}b^{1/2}n^{3/4})$. Thus, removing the vertical and horizontal lines does not affect our incidence bound. Note that every non-horizontal line contains at most a points of $A \times B$, so we now have $I(A \times B, \mathcal{L}) \leq an$. If $an = O(a^{3/4}b^{1/2}n^{3/4})$ then we are done. We may thus assume that $an = \Omega(a^{3/4}b^{1/2}n^{3/4})$, which implies that $b = O(\sqrt{an})$.

In our second pruning step, we check whether $\mathcal{L}$ contains a subset of more than $\sqrt{an}$ lines that are all parallel or concurrent. If such a subset exists, then we remove all of its lines from $\mathcal{L}$ and repeat this process. Let n_j be the number of lines that were removed during the j'th iteration of this process. Excluding the point of concurrency, every point of $A \times B$ is incident to at most one line that was removed in the j'th iteration. That is, the number of incidences that are removed in the j'th iteration is at most $n_j + ab$. Since the number of iterations is smaller than $n/\sqrt{an} = \sqrt{n/a}$, the number of incidences that are removed during the second pruning step is smaller than $\sum_j (n_j + ab) = n + \sqrt{n/a} \cdot ab = n + b\sqrt{na}$. Since $b = O(\sqrt{an})$, the number of removed incidences is $O(n + b^{1/2}a^{3/4}n^{3/4})$.

We consider the dual set

$$\mathcal{L}^* = \{(s, t) : \ \mathcal{L} \text{ contains a line defined by } y = sx + t\}.$$

Since $\mathcal{L}$ does not contain vertical lines, every line of $\mathcal{L}$ has a dual point in $\mathcal{L}^*$. For $\beta \in B$, we set

$$r_\beta = |\{(\alpha, s, t) \in A \times \mathcal{L}^* : \ \beta = \alpha s + t\}|.$$

Intuitively, r_β is the number of incidences that occur on the line defined by $y \equiv \beta$. We thus have that $I(A \times B, \mathcal{L}) = \sum_{\beta \in B} r_\beta$. We also set the energy variant

$$E = \left|\left\{(\alpha, s, t, \alpha', s', t') \in (A \times \mathcal{L}^*)^2 : \ \alpha s + t = \alpha' s' + t'\right\}\right|.$$

We complete the proof by double counting E. The Cauchy–Schwarz inequality implies

$$E = \sum_{\beta \in B} r_\beta^2 \geq \frac{\left(\sum_{\beta \in B} r_\beta\right)^2}{b} = \frac{I(A \times B, \mathcal{L})^2}{b}. \tag{13.9}$$

We refer to the coordinates of $\mathbb{F}_q^3$ as x, y, z. We reduce the problem of deriving an upper bound for E to a point-plane incidence problem in $\mathbb{F}_q^3$, as follows. We define the point set

$$\mathcal{Q} = \{(\alpha, s', t') \in A \times \mathcal{L}^*\},$$

and the set of planes

$$\Pi = \{xs + t = \alpha'y + z : \ (s, t) \in \mathcal{L}^* \text{ and } \alpha' \in A\}.$$

Note that $E = I(\mathcal{Q}, \Pi)$. To apply Theorem 13.6 with $\mathcal{Q}$ and Π, we check that the conditions of this theorem are satisfied. We indeed have $|\mathcal{Q}| = a|\mathcal{L}| = |\Pi|$ and $|\mathcal{Q}| = a|\mathcal{L}| = O(p^2)$. It remains to obtain an upper bound for the number of points of $\mathcal{Q}$ that can be on a common line. If a set of points from $\mathcal{L}^* \subset \mathbb{F}_q^2$ are on a common line defined by $y = cx + d$, then the corresponding lines of $\mathcal{L}$ intersect in the point $(-c, d) \in \mathbb{F}_q^2$. By the second pruning step of $\mathcal{L}$, no line in $\mathbb{F}_q^2$ contains $\sqrt{an}$ points of $\mathcal{L}^*$. Thus, any line in $\mathbb{F}_q^3$ that is not parallel to the x-axis contains fewer than $\sqrt{an}$ points of $\mathcal{Q}$. A line that is parallel to the x-axis contains at most a points of $\mathcal{Q}$, and by an assumptions of this theorem $a = O(\sqrt{an})$. We conclude that every line in $\mathbb{F}_q^3$ contains $O(\sqrt{an})$ points of $\mathcal{Q}$.

By the above, we may apply Theorem 13.6 on $\mathcal{Q}$ and Π, with $k = O(\sqrt{an})$. This implies

$$E = I(\mathcal{Q}, \Pi) = O(|\Pi|\sqrt{|\mathcal{Q}|} + k|\Pi|) = O(a^{3/2} n^{3/2}).$$

Combining this with (13.9) immediately implies the assertion of the theorem. $\square$

As in the proof of Theorem 13.11, establishing an improved bound on k in the above proof does not lead to new results.

Our next goal is to rely on the above bound for incidences with a lattice to obtain a bound for the general incidence problem. The following lemma shows that when there are many point-line incidences in $\mathbb{F}_q^2$, a large portion of the point set behaves like a lattice.

Below we prove Theorem 13.12 by induction. As discussed in Chapter 7, using $O(\cdot)$-notation in such proofs can be problematic. We thus avoid using $O(\cdot)$-notation starting now.

Lemma 13.14. *The following holds for any real constants $0 < c_1 < c_2$. Let $\mathcal{P}$ be a set of m points and let $\mathcal{L}$ be a set of n lines, both in $\mathbb{F}_q^2$. Let*

$$r \geq \max \left\{ 4n/c_1 m, 4/c_1, (2^5 n^2 / c_1^3 m)^{1/3} \right\}$$

satisfy that every point of $\mathcal{P}$ is incident to at least $c_1 r$ lines of $\mathcal{L}$ and to at most $c_2 r$ such lines. Then there exist distinct points $u, v \in \mathbb{F}_q^2$, a subset $\mathcal{P}' \subset \mathcal{P}$, and line sets $\mathcal{L}_u, \mathcal{L}_v \subset \mathcal{L}$, with the following properties. The set $\mathcal{P}'$ contains no points that are on the line incident to both u and v, and $|\mathcal{P}'| \geq m \frac{c_1^4 r^4}{2^7 n^2}$. Every point of $\mathcal{P}'$ is incident to a line of $\mathcal{L}_u$ and to a line of $\mathcal{L}_v$, and each of these sets contains at most $c_2 r$ lines.

To see why the set $\mathcal{P}'$ from Lemma 13.14 behaves like a lattice, consider this set in the projective plane $\mathbb{PF}_q^3$ by giving each point a z-coordinate of 1 (see Section 13.2). As will be explained in the proof of Theorem 13.12 below, there exists a projective transformation that takes the lines of $\mathcal{L}_u$ to lines parallel to the x-axis and the lines of $\mathcal{L}_v$ to lines parallel to the y-axis. These two sets of lines define a $(c_2 r) \times (c_2 r)$ lattice that contains $\mathcal{P}'$.

Proof of Lemma 13.14. Set $x = I(\mathcal{P}, \mathcal{L})$, and let $\mathcal{L}_+$ be the set of lines of $\mathcal{L}$ that are incident to at least $x/2n$ points of $\mathcal{P}$. We have

$$I(\mathcal{P}, \mathcal{L}_+) = I(\mathcal{P}, \mathcal{L}) - I(\mathcal{P}, \mathcal{L} \setminus \mathcal{L}_+) \geq x - n\frac{x}{2n} = x/2. \tag{13.10}$$

If every point of $\mathcal{P}$ is incident to fewer than $x/2m$ lines of $\mathcal{L}_+$ then $I(\mathcal{P}, \mathcal{L}_+) < m\frac{x}{2m} = x/2$. Since this contradicts (13.10), there exists a point $u \in \mathcal{P}$ that is incident to at least $x/2m$ lines of $\mathcal{L}_+$. Since every point of $\mathcal{P}$ is incident to at least $c_1 r$ lines of $\mathcal{L}$, we have $x \geq mc_1 r$. That is, u is incident to at least $c_1 r/2$ lines of $\mathcal{L}_+$. We set

$$\hat{\mathcal{P}} = \{v \in \mathcal{P} \setminus \{u\} : \ u, v \in \ell \text{ for some } \ell \in \mathcal{L}\},$$

and $\hat{m} = |\hat{\mathcal{P}}|$. Every line of $\mathcal{L}_+$ is incident to at least $x/2n \geq mc_1r/2n$ points of $\mathcal{P}$. Combining this with the assumption $r \geq 4n/c_1m$ leads to

$$\hat{m} \geq \frac{c_1r}{2}\left(\frac{mc_1r}{2n} - 1\right) \geq \frac{mc_1^2r^2}{8n}. \tag{13.11}$$

We now repeat the above analysis for the points of $\hat{\mathcal{P}}$. Set $\hat{x} = I(\hat{\mathcal{P}}, \mathcal{L})$, and let $\hat{\mathcal{L}}_+$ be the set of lines of $\mathcal{L}$ that are incident to at least $\hat{x}/2n$ points of $\hat{\mathcal{P}}$. We have

$$I(\hat{\mathcal{P}}, \hat{\mathcal{L}}_+) = I(\hat{\mathcal{P}}, \mathcal{L}) - I(\hat{\mathcal{P}}, \mathcal{L} \setminus \hat{\mathcal{L}}_+) \geq \hat{x} - n\frac{\hat{x}}{2n} = \hat{x}/2. \tag{13.12}$$

If every point of $\hat{\mathcal{P}}$ is incident to fewer than $\hat{x}/2\hat{m}$ lines of $\hat{\mathcal{L}}_+$ then $I(\hat{\mathcal{P}}, \hat{\mathcal{L}}_+) < m'\frac{\hat{x}}{2\hat{m}} = \hat{x}/2$. Since this contradicts (13.12), there exists a point $v \in \hat{\mathcal{P}}$ that is incident to at least $\hat{x}/2\hat{m}$ lines of $\hat{\mathcal{L}}_+$. Since every point of $\hat{\mathcal{P}}$ is incident to at least c_1r lines of $\mathcal{L}$, we have $\hat{x} \geq \hat{m}c_1r$. That is, u is incident to at least $c_1r/2$ lines of $\hat{\mathcal{L}}_+$. Let ℓ_{uv} be the line incident to u and v. We set

$$\mathcal{P}' = \{w \in \hat{\mathcal{P}} \setminus \ell_{uv} : \ w, v \in \hat{\ell} \text{ for some } \hat{\ell} \in \hat{\mathcal{L}}\}.$$

Every line of $\hat{\mathcal{L}}_+$ is incident to at least $\hat{x}/2n \geq \hat{m}c_1r/2n$ points of $\mathcal{P}$. Combining this with (13.11) and with the assumption $r \geq \max\left\{4/c_1, (2^5n^2/c_1^3m)^{1/3}\right\}$ leads to

$$|\mathcal{P}'| \geq \left(\frac{c_1r}{2} - 1\right)\left(\frac{\hat{m}c_1r}{2n} - 1\right) \geq \frac{mc_1^4r^4}{2^7n}.$$

Let $\mathcal{L}_u$ be the lines of $\mathcal{L} \setminus \{\ell_{uv}\}$ that are incident to u, and similarly for $\mathcal{L}_v$. By the assumption involving c_2, we have that $|\mathcal{L}_u| \leq c_2r$ and $|\mathcal{L}_v| \leq c_2r$. By definition, every point of $\mathcal{P}'$ is incident to a line of $\mathcal{L}_v$. Since every point of $\hat{\mathcal{P}}$ is incident to a line of $\mathcal{L}_u$, so is every point of $\mathcal{P}' \subset \hat{\mathcal{P}}$. $\qquad\square$

We combine Theorem 13.13 and Lemma 13.14 to obtain a general point-line incidence bound.

Proof of Theorem 13.12. In the following we assume that $m^{7/8} < n < m^{8/7}$ and $n^{13}/m^2 = O(p^{15})$, as assumed in the theorem. We will prove that there exists a sufficiently large constant c such that for any set $\mathcal{P}$ of m and a set $\mathcal{L}$ of n lines, both in $\mathbb{F}_q^2$,

$$I(\mathcal{P}, \mathcal{L}) < cm^{11/15}n^{11/15}. \tag{13.13}$$

For small values of n, we get that (13.13) holds by taking c to be sufficiently large. We may thus assume that n is larger than some constant (say, $n \geq 100$). For any

such fixed value of n, we prove (13.13) by induction on m. For the induction basis, note that (13.1) implies (13.13) when $n^{4/11} \leq m \leq n^{7/8}$ and c is sufficiently large.

To prove the induction step, we consider $m > n^{7/8}$ and assume that (13.13) holds for any smaller m that satisfies $m \geq n^{4/11}$. We assume for contradiction that there exists a set of m points $\mathcal{P}$ and a set of n lines $\mathcal{L}$, both in $\mathbb{F}_q^2$, such that $I(\mathcal{P}, \mathcal{L}) \geq cm^{11/15}n^{11/15}$. If $I(\mathcal{P}, \mathcal{L}) > 2cm^{11/15}n^{11/15}$ then we can decrease the number of incidences by moving some points of $\mathcal{P}$, to obtain $cm^{11/15}n^{11/15} \leq I(\mathcal{P}, \mathcal{L}) \leq 2cm^{11/15}n^{11/15}$. We may thus assume that $I(\mathcal{P}, \mathcal{L}) \leq 2cm^{11/15}n^{11/15}$. Let $r = I(\mathcal{P}, \mathcal{L})/m$, and consider the sets

$$\mathcal{P}_{\text{poor}} = \{p \in P : p \text{ is incident to at most } 2^{-2}r \text{ lines of } \mathcal{L}\},$$
$$\mathcal{P}_{\text{rich}} = \{p \in P : p \text{ is incident to at least } 2^3 r \text{ lines of } \mathcal{L}\},$$
$$\mathcal{P}_{\text{mid}} = \mathcal{P} \setminus (\mathcal{P}_{\text{poor}} \cup \mathcal{P}_{\text{rich}}).$$

Note that $I(\mathcal{P}_{\text{poor}}) \leq m \cdot 2^{-2}r = 2^{-2}I(\mathcal{P}, \mathcal{L})$. We also observe that $|\mathcal{P}_{\text{rich}}| \leq I(\mathcal{P}, \mathcal{L})/(2^3 r) = 2^{-3}m$. If we also have that $|\mathcal{P}_{\text{rich}}| \geq n^{4/11}$, then by the induction hypothesis $I(\mathcal{P}_{\text{rich}}, \mathcal{L}) < 2^{-2}I(\mathcal{P}, \mathcal{L})$. If $|\mathcal{P}_{\text{rich}}| < n^{4/11}$, then (13.1) and a sufficiently large c imply $I(\mathcal{P}_{\text{rich}}, \mathcal{L}) = O(n) < 2^{-2}I(\mathcal{P}, \mathcal{L})$.

If $|\mathcal{P}_{\text{mid}}| \leq 2^{-3}m$ then by applying the induction hypothesis to $\mathcal{P}_{\text{mid}}$ and $\mathcal{L}$ we get $I(\mathcal{P}_{\text{mid}}, \mathcal{L}) < 2^{-2}I(\mathcal{P}, \mathcal{L})$ (as in the previous paragraph, if $|\mathcal{P}_{\text{mid}}| < n^{4/11}$ then we use (13.1) instead of the hypothesis). This leads to the contradiction $I(\mathcal{P}, \mathcal{L}) = I(\mathcal{P}_{\text{poor}}, \mathcal{L}) + I(\mathcal{P}_{\text{rich}}, \mathcal{L}) + I(\mathcal{P}_{\text{mid}}, \mathcal{L}) < 3I(\mathcal{P}, \mathcal{L})/4$, so we may assume that $|\mathcal{P}_{\text{mid}}| > 2^{-3}m$. We set $\mathcal{P}_1 = \mathcal{P}_{\text{mid}}$, $m_1 = |\mathcal{P}_1|$, $c_1 = 2^{-2}$, and $c_2 = 2^3$. To be able to apply Lemma 13.14, we claim that

$$\frac{I(\mathcal{P}, \mathcal{L})}{m} = r \geq \max \left\{ \frac{4n}{c_1 m_1}, \frac{4}{c_1}, \left(\frac{2^5 n^2}{c_1^3 m_1} \right)^{1/3} \right\}. \tag{13.14}$$

It is not difficult to verify that (13.14) holds for sufficiently large c, since $I(\mathcal{P}, \mathcal{L}) \geq cm^{11/15}n^{1/15}$, $m_1 \geq 2^{-3}m$, and $m^{7/8} < n < m^{8/7}$.

By (13.14) and the definition of $\mathcal{P}_{\text{mid}}$, we can apply Lemma 13.14 on $\mathcal{P}_1$ and $\mathcal{L}$, with the above values for c_1, c_2, r. We obtain points $u_1, v_1 \in \mathcal{P}_1$, a set $\mathcal{P}_1' \subset \mathcal{P}_1$, and sets of lines $\mathcal{L}_{u,1}, \mathcal{L}_{v,1} \subset \mathcal{L}$, that satisfy the assertions of Lemma 13.14. We set $\mathcal{P}_2 = \mathcal{P}_1 \setminus \mathcal{P}_1'$. We repeat the above process as follows. At the j'th iteration we apply Lemma 13.14 on $\mathcal{P}_j$ and $\mathcal{L}$ with the above values for c_1, c_2, and r. We obtain points $u_j, v_j \in \mathcal{P}_{\text{mid}}$, a set $\mathcal{P}_j' \subset \mathcal{P}_{\text{mid}}$, and sets of lines $\mathcal{L}_{u,j}, \mathcal{L}_{v,j} \subset \mathcal{L}$. We then set $\mathcal{P}_{j+1} = \mathcal{P}_j \setminus \mathcal{P}_j'$. The process stops once get that $|\mathcal{P}_{j+1}| \leq 2^{-3}m$. As long as $|\mathcal{P}_{j+1}| > 2^{-3}m$, the conditions of (13.14) remain valid and we may keep applying Lemma 13.14.

Let s be the number of times we applied Lemma 13.14 in the above process. By that lemma, for every $1 \le j \le s$ we have

$$|\mathcal{P}_j'| \ge |\mathcal{P}_j|\frac{c_1^4 r^4}{2^7 n^2} \ge 2^{-3} m \frac{c_1^4 r^4}{2^7 n^2} \ge 2^{-18} m \frac{r^4}{n^2}.$$

We can thus bound the number of iterations by

$$s \le \frac{m}{2^{-18} m r^4/n^2} = \frac{n^2 2^{18}}{r^4}. \tag{13.15}$$

We now consider $I(\mathcal{P}_j, \mathcal{L})$ for some $1 \le j \le s$. By Lemma 13.14, every point of $\mathcal{P}_j$ is incident to a line of $\mathcal{L}_{u,j}$ and to a line of $\mathcal{L}_{v,j}$. We move from the affine plane $\mathbb{F}_q^2$ to the projective plane $\mathbb{PF}_q^2$, as explained in Section 13.2. We then perform a projective transformation τ_j that takes u_j to $(1,0,0)$ and v_j to $(0,1,0)$. By inspecting the definition in (13.4), we observe that such a transformation always exists.[4] As explained in Section 13.2, τ_j takes every line of $\mathcal{L}_{u,j}$ to a line that is parallel to the y-axis, and every line of $\mathcal{L}_{v,j}$ to a line parallel to the x-axis. This in turn implies that τ_j takes $\mathcal{P}_j$ to a subset of an $(c_2 r) \times (c_2 r)$ lattice G_j. We then return to the affine plane $\mathbb{F}_q^2$ by taking the intersection of $\mathbb{PF}_q^3$ with the plane defined by $z \equiv 1$. By the statement of Lemma 13.14, no point of $\mathcal{P}_j$ is on the line incident to both u_j and v_j, so τ_j does not take any point of $\mathcal{P}_j$ to the line at infinity. Thus, when returning to the affine plane $\mathbb{F}_q^2$ we do not lose any incidences between $\mathcal{P}_j$ and $\mathcal{L}$.

Let α be the constant hidden in the $O(\cdot)$-notation in the bound of Theorem 13.13. We now verify that the condition of Theorem 13.13 hold for G_j and a set of n lines in $\mathbb{F}_q^2$. By assuming that n is sufficiently large, we indeed have that $(c_2 r)^3 = O(n)$. Combining $I(\mathcal{P}, \mathcal{L}) \le 2cm^{11/15}n^{11/15}$ with $n^{13}/m^2 = O(p^{15})$ implies that $c_2 rn = O(p^2)$. Since projective transformations preserve incidences, Theorem 13.13 yields

$$I(\mathcal{P}_j, \mathcal{L}) = I(\tau_j(\mathcal{P}_j), \tau_j(\mathcal{L})) \le I(G_j, \tau_j(\mathcal{L})) \le \alpha \left((c_2 r)^{3/4}(c_2 r)^{1/2} n^{3/4} + n \right).$$

Combining this with (13.15) and assuming that c is sufficiently large gives

$$\sum_{j=1}^{s} I(\mathcal{P}_j, \mathcal{L}) \le \frac{n^2 2^{18}}{r^4} \cdot \alpha \left((c_2 r)^{5/4} n^{3/4} + n \right) < \alpha \left(2^{22} \frac{n^{11/4}}{r^{11/4}} + \frac{n^3 2^{18}}{r^4} \right)$$

$$= \alpha \left(2^{22} \frac{n^{11/4} m^{11/4}}{I(\mathcal{P}, \mathcal{L})^{11/4}} + \frac{n^3 m^4 2^{18}}{I(\mathcal{P}, \mathcal{L})^4} \right) < 2^{-2} I(\mathcal{P}, \mathcal{L}).$$

---

[4]For some intuition, note that in $\mathbb{R}^2$ we can take any pair of points to $(1,0)$ and $(0,1)$ by scaling, translating, and rotating the plane.

Since $|\mathcal{P}_{s+1}| \leq 2^{-3}m$, by the induction hypothesis $I(\mathcal{P}_{s+1}, \mathcal{L}) < 2^{-2}I(\mathcal{P}, \mathcal{L})$ (as before, if $|\mathcal{P}_{s+1}| < n^{4/11}$ then we use (13.1) instead of the hypothesis). Combining all of the above cases gives

$$I(\mathcal{P}, \mathcal{L}) = I(\mathcal{P}_{\text{poor}}, \mathcal{L}) + I(\mathcal{P}_{\text{rich}}, \mathcal{L}) + \sum_{j=1}^{s+1} I(\mathcal{P}_j, \mathcal{L}) < I(\mathcal{P}, \mathcal{L}).$$

This contradiction completes the induction step, and thus the proof of the theorem. $\square$

13.7 Exercises

Problem 13.1. Let $\mathcal{P}$ be a set of m points in $\mathbb{R}^3$. Let $\mathcal{V}$ be a set of n sphere in $\mathbb{R}^3$, all of radius one and centered in points with a zero z-coordinate. Use duality to prove that $I(\mathcal{P}, \mathcal{V}) = O(n^{3/5}m^{4/5} + m + n)$.

Problem 13.2. Prove derive the sum-product bound in Theorem 13.11 by using the point–line incidence bound of Theorem 13.12 (rather than the point–plane incidence bound of Theorem 13.6).

Problem 13.3. For $f(x, y) = x^2 + xy$ and $A \subset \mathbb{F}_q$, set $f(A, A) = \{f(a, b) : a, b \in A\}$. Show that when $|A| = O(p^{2/3})$ we have $|f(A, A)| = \Omega(|A|^{4/5})$. (Hint: Use energy.)

Problem 13.4. Let $A \subset \mathbb{F}_q$ with $|A| = O(|A|^{2/3})$. Let $A + AA = \{a + bc : a, b, c \in A\}$. Adapt the use of E in the proof of Theorem 13.13 to show that $|A + AA| = \Omega(|A|^{3/2})$. (This is one case where we have similar results over and over $\mathbb{F}_q$.)

13.8 Open problems

Let $\mathcal{P}$ be a set of m lines and let $\mathcal{L}$ be a set of n lines, both in $\mathbb{F}_q^3$. In this chapter we saw that the Szemerédi–Trotter bound $I(\mathcal{P}, \mathcal{L}) = O(|\mathcal{P}|^{2/3}|\mathcal{L}|^{2/3} + |\mathcal{P}| + |\mathcal{L}|)$ does not hold when mn is asymptotically larger than q^3. We saw that this bound does hold when $mn = \Theta(q^3)$, and also in some cases where m and n are tiny compared to q (about $\log \log \log q$). There is a huge gap between these two extreme cases, and we are far from understanding what happens in it. The current best bound, stated in Theorem 13.12, might be very far from begin tight.

Recall that the $o(\cdot)$-notation means "asymptotically smaller than".

Open Problem 13.1. Let $\mathcal{P}$ be a set of m points and let $\mathcal{L}$ be a set of n lines, both in $\mathbb{F}_q^3$, such that $mn = o(q^3)$. Find a tight upper bound for $I(\mathcal{P}, \mathcal{L})$.

The bound of Theorem 13.6 is sharp when $k = \Omega(\sqrt{m})$. When $k = o(\sqrt{m})$, it seems plausible that a better bound holds. Such a bound does exist in $\mathbb{R}^3$.

Open Problem 13.2. Let $q = p^r$ for some prime p and positive integer r. Let $\mathcal{P}$ be a set of m points and let Π be a set of n planes, both in $\mathbb{F}_q^3$, such that $n \geq m$. Assume that $m = O(p^2)$ and that no line of $\mathbb{F}_q^3$ contains k points of $\mathcal{P}$, for some $k = o(r)$. Find a tight upper bound for $I(\mathcal{P}, \Pi)$.

Bibliography

[1] M. Ajtai, V. Chvátal, M. M. Newborn, and E. Szemerédi, Crossing-free subgraphs, *Annals Discrete Math.* **12** (1982), 9–12.

[2] N. Alon, J. H. Spencer, *The Probabilistic Method*, Wiley–Interscience, 2004.

[3] M. F. Atiyah, *Collected works. Vol. 6*, Oxford Science Publications, The Clarendon Press Oxford University Press, 2004.

[4] S. Bardwell-Evans and A. Sheffer, A Reduction for the Distinct Distances Problem in $\mathbb{R}^d$, arXiv:1705.10963.

[5] S. Barone, and S. Basu, Refined bounds on the number of connected components of sign conditions on a variety, *Discrete Comput. Geom.* **47** (2012), 577–597.

[6] S. Basu and O. E. Raz, An o-minimal Szemerédi-Trotter theorem, arXiv:1611.07362.

[7] S. Basu and M. Sombra, Polynomial partitioning on varieties of codimension two and point-hypersurface incidences in four dimensions, *Discrete Comput. Geom.* **55** (2016), 158–184.

[8] M. Bateman and N. Katz, New bounds on cap sets, *Journal of the American Mathematical Society* **25** (2012), 585–613.

[9] J. Beck, On the lattice property of the plane and some problems of Dirac, Motzkin and Erdos in combinatorial geometry, *Combinatorica* **3** (1983), 281–297.

[10] F. A. Behrend, On sets of integers which contain no three terms in arithmetical progression, *Proceedings of the National Academy of Sciences of the United States of America* **32** (1946), 331.

[11] B. C. Berndt and R. A. Rankin, *Ramanujan: Letters and Commentary*, Amer. Math. Soc., Providence, RI, 1995.

[12] J. Bochnak, M. Coste, and M. Roy, *Real Algebraic Geometry*, Springer-Verlag, Berlin, 1998.

[13] E. Bombieri and J. Pila, The number of integral points on arcs and ovals, *Duke Math. J.* **59** (1989), 337–357.

[14] M. Bóna, A Walk Through Combinatorics, World Scientific, Singapore, 2002.

[15] J. Bourgain and C. Demeter, New Bounds for the Discrete Fourier Restriction to the Sphere in 4D and 5D, *Int. Math. Res. Not.* (2015), 3150–3184.

[16] C. B. Boyer, The Invention of Analytic Geometry, *Scientific American* (1949).

[17] E. Breuillard, B. Green, and T. Tao, Approximate subgroups of linear groups, *Geometric and Functional Analysis* **21** (2011), 774–819.

[18] M. Charalambides, A note on distinct distance subsets, *Journal of Geometry*, **104** (2013), 439–442.

[19] B. Chazelle, H. Edelsbrunner, L. J. Guibas, R. Pollack, R. Seidel, M. Sharir, and J. Snoeyink, Counting and cutting cycles of lines and rods in space, *Computational Geometry* 1 (1992), 305–323.

[20] C. Chen, B. Kerr, and A. Mohammadi, A new sum–product estimate in prime fields, *Bulletin of the Australian Mathematical Society* (2018), 1–13.

[21] A. Chernikov, D. Galvin, and S. Starchenko, Cutting lemma and Zarankiewicz's problem in distal structures, arXiv:1612.00908.

[22] K. L. Clarkson, H. Edelsbrunner, L. J. Guibas, M. Sharir, and E. Welzl, Combinatorial complexity bounds for arrangements of curves and spheres, *Discrete Comput. Geom.* **5** (1990), 99–160.

[23] D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*, 3rd edition, Springer-Verlag, Heidelberg, 2007.

[24] E. Croot, V. F. Lev, and P. P. Pach, Progression-free sets in are exponentially small, *Annals of Mathematics 185* (2017), 331–337.

[25] C. Demeter, Incidence theory and restriction estimates, arXiv:1401.1873.

[26] Z. Dvir, On the size of Kakeya sets in finite fields, *Journal of the American Mathematical Society* **22** (2009), 1093–1097.

[27] Z. Dvir, Incidence theorems and their applications, *Foundations and Trends in Theoretical Computer Science* **6** (2012), 257–393.

[28] Z. Dvir and S. Gopi, On the number of rich lines in truly high dimensional sets, *Proc. 31st ACM Symp. on Computational Geometry* (2015), 584–598

[29] Y. Edel, Extensions of generalized product caps, *Designs, Codes and Cryptography* **31** (2004), 5–14.

[30] G. Elekes, On the number of sums and products, *ACTA ARITHMETICA — WARSZAWA* **81** (1997), 365–367.

[31] G. Elekes, A note on the number of distinct distances, *Period. Math. Hung.*, **38** (1999), 173–177.

[32] G. Elekes, Sums versus products in Number Theory, Algebra and Erdős Geometry — a survey, in *Paul Erdős and his Mathematics II*, Bolyai Math. Soc. Stud 11 Budapest, 2002, 241–290.

[33] G. Elekes and E. Szabó, How to find groups?(and how to use them in Erdős geometry?), *Combinatorica* **32** (2012), 537–571.

[34] J. S. Ellenberg and D. Gijswijt, On large subsets of with no three-term arithmetic progression, *Annals of Mathematics* **185** (2017), 339–343.

[35] P. Erdős, On sets of distances of n points, *Amer. Math. Monthly* **53** (1946), 248–250.

[36] P. Erdős, Problems and results in combinatorial geometry, Discrete geometry and convexity, *Annals of New York Acad. Sci.* **440** (1985), 1–11.

[37] P. Erdős, On some metric and combinatorial geometric problems, *Discrete Math.* **60** (1986), 147–153.

[38] P. Erdős, On some of my favourite theorems, *Combinatorics, Paul Erdős is Eighty,* Vol. 2 (D. Miklós et al., eds.), Bolyai Society Mathematical Studies 2, Budapest, 1996, 97–132.

[39] P. Erdős and G. Purdy, Some extremal problems in geometry, *J. Combinat. Theory* **10** (1971), 246–252.

[40] P. Erdős and E. Szemerédi, On sums and products of integers, *Studies in pure mathematics*. Birkhäuser Basel, 1983, 213–218.

[41] M. Farber, S. Ray, and S. Smorodinsky, On totally positive matrices and geometric incidences, *J. Combinat. Theory A* **128** (2014), 149–161.

[42] J. Fox, J. Pach, A. Sheffer, A. Suk, and J. Zahl, A semi-algebraic version of Zarankiewicz's problem, *J. of the European Mathematical Society*, **19** (2017), 1785–1810.

[43] J. Fox, J. Pach, and A. Suk, More distinct distances under local conditions, *Combinatorica*, to appear.

[44] M. Giusti, Some effectivity problems in polynomial ideal theory, in *Proc. EU-ROSAM 84*, Lect. Notes in Comp. Sci., **174** (1984), 159–171.

[45] B. Green and T. Tao, On sets defining few ordinary lines, *Discrete Comput. Geom.* **50** (2013), 409–468.

[46] E. Grosswald, *Representations of integers as sums of squares*, Springer Science & Business Media, 2012.

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

[48] L. Guth, Distinct distance estimates and low degree polynomial partitioning, *Discrete Comput. Geom.* **53** (2015), 428–444.

[49] L. Guth, *Polynomial Methods in Combinatorics*, American Mathematical Soc., 2016.

[50] L. Guth, and N. H. Katz, Algebraic methods in discrete analogs of the Kakeya problem, *Advances in Mathematics* **225** (2010), 2828–2839.

[51] L. Guth and N.H. Katz, On the Erdős distinct distances problem in the plane, *Annals Math.* **181** (2015), 155–190.

[52] M. Hablicsek and Z. Scherr, On the number of rich lines in high dimensional real vector spaces, *Discrete Comput. Geom.* **55** (2016), 955–962.

[53] P. R. Halmos, How to write mathematics, *Enseign. Math* **16** (1970), 123–152.

[54] J. Harris, *Algebraic geometry: a first course*, Springer, New York, 1992.

[55] R. Hartshorne, *Algebraic Geometry*, Springer-Verlag, New York. 1983.

[56] A. Iosevich, Fourier analysis and geometric combinatorics, In *Topics in mathematical analysis*, 321–335, World Sci. Publ., Hackensack, NJ, 2008.

[57] H. Kaplan, M. Sharir, and E. Shustin, On lines and joints, *Discrete Comput. Geom.* **44** (2010), 838–843.

[58] W. G. Kelley and A. C. Peterson, *The theory of differential equations: classical and qualitative*, Springer Science & Business Media, 2010.

[59] J. Kollár, Szemerédi–Trotter-type theorems in dimension 3, *Advances in Mathematics* **271** (2015), 30–61.

[60] E. Landau, Über die Einteilung der positiven ganzen Zahlen in vier Klassen nach der Mindeszahl der zu ihrer additiven Zusammensetzung erforderlichen Quadrate, *Arch. Math. Phys.* **13** (1908), 305–312.

[61] D. G. Larman, C. A. Rogers, and J. J. Seidel, On two-distance sets in euclidean space *Bulletin of the London Mathematical Society* **9** (1977), 261–267.

[62] F. T. Leighton, *Complexity issues in VLSI: optimal layouts for the shuffle-exchange graph and other networks*, MIT press, 1983.

[63] C. Y. Lo, J. Matoušek, and W. L. Steiger, Algorithms for ham-sandwich cuts, *Discrete Comput. Geom.* **11** (1994), 433–452.

[64] J. Matoušek, *Lectures on Discrete Geometry*, Springer Verlag, Heidelberg, 2002.

[65] J. Matoušek, The number of unit distances is almost linear for most norms, *Advances in Mathematics* **226** (2011), 2618–2628.

[66] R. Meshulam, On subsets of finite abelian groups with no 3-term arithmetic progressions, *Journal of Combinatorial Theory A* **71** (1995), 168–172.

[67] J. Milnor, On the Betti numbers of real varieties, *Proc. Amer. Math. Soc.* **15** (1964), 275–280.

[68] L. Moser, On the different distances determined by n points, *American Mathematical Monthly* **59** (1952), 85–91.

[69] J. Pach and M. Sharir, Repeated angles in the plane and related problems, *J. Combinat. Theory Ser. A* **59** (1992), 12–22.

[70] J. Pach and M. Sharir, On the number of incidences between points and curves. *Combinat. Probab. Comput.* **7** (1998), 121–127.

[71] J. Pach and G. Tardos, Isosceles triangles determined by a planar point set, *Graphs and Combinatorics* **18** (2002), 769–779.

[72] J. Pach and F. de Zeeuw, Distinct distances on algebraic curves in the plane, *Comb. Probab. Comp* **26** (2017): 99–117.

[73] R. Quilodrán, The Joints Problem in $\mathbb{R}^n$, *SIAM Journal on Discrete Math.* **23** (2010), 2211–2213.

[74] O. E. Raz and M. Sharir, Unit-area triangles in the plane: Theme and variations, *Combinatorica*, to appear.

[75] O. E. Raz, M. Sharir, and J. Solymosi, Polynomials vanishing on grids: The Elekes-Rónyai problem revisited, *Amer. J. Math.* **138** (2016), 1029–1065.

[76] O. E. Raz, M. Sharir, and F. De Zeeuw, Polynomials vanishing on Cartesian products: The Elekes–Szabó theorem revisited, *Duke Mathematical Journal* **165** (2016): 3517–3566.

[77] C. Reid, *Hilbert: with an appreciation of Hilbert's mathematical work by Hermann Weyl*, Springer, 1970.

[78] O. Roche-Newton, M. Rudnev, and I. D. Shkredov, New sum-product type estimates over finite fields, *Advances in Mathematics* **293** (2016), 589–605.

[79] G.C. Rota, *Indiscrete Thoughts,* Birkhäuser Boston, 1997.

[80] K. F. Roth, On certain sets of integers, *J. London Math. Soc.* **28** (1953), 245–252.

[81] M. Rudnev, On the number of incidences between planes and points in three dimensions, *Combinatorica* **38** (2018): 219–254.

[82] J. T. Schwartz, Fast probabilistic algorithms for verification of polynomial identities, *J. ACM* **27** (1980), 701–717.

[83] M. Sharir and J. Zahl, Cutting algebraic curves into pseudo-segments and applications, *J. Combinat. Theory Ser. A*, to appear.

[84] A. Sheffer, Lower bounds for incidences with hypersurfaces, *Discrete Analysis* 2016:16.

[85] A. Sheffer, Distinct Distances: Open Problems and Current Bounds, arXiv:1406.1949.

[86] M. Sharir, A. Sheffer, and N. Solomon, Incidences with curves in $\mathbb{R}^d$, *Electronic J. Combinat.* **23** (2016), P4.16.

[87] M. Sharir, A. Sheffer, and J. Solymosi, Distinct distances on two lines, *J. Combinat. Theory Ser. A*, **120** (2013), 1732–1736.

[88] A. Sheffer, E. Szabó, and J. Zahl, Point-curve incidences in the complex plane, *Combinatorica*, to appear.

[89] J. Solymosi, Bounding multiplicative energy by the sumset, *Advances in mathematics* **222** (2009): 402–408.

[90] J. Solymosi and T. Tao, An incidence theorem in higher dimensions, *Discrete Comput. Geom.* **48** (2012), 255–280.

[91] J. Solymosi and V. Vu, Near optimal bounds for the Erdős distinct distances problem in high dimensions, *Combinatorica* **28** (2008), 113–125.

[92] J. Spencer, E. Szemerédi, and W. T. Trotter, Unit distances in the Euclidean plane, *Graph Theory and Combinatorics* (ed. B. Bollobás), Academic Press, 1984.

[93] S. Stevens and F. de Zeeuw, An improved point-line incidence bound over arbitrary fields, *Bulletin of the London Mathematical Society* **49** (2017), 842–858.

[94] J. Stillwell, *Naive lie theory*, Springer, New York, London, 2008.

[95] A. H. Stone and J. W. Tukey, Generalized sandwich theorems, *Duke Math. J.* **9** (1942), 356–359.

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

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

[98] T. Tao, A symmetric formulation of the-Croot-Lev-Pach-Ellenberg-Gijswijt capset bound, https://terrytao.wordpress.com/2016/05/18/a-symmetric-formulation-of-the-croot-lev-pach-ellenberg-gijswijt-capset-bound/

[99] T. Tao and V. H. Vu, *Additive combinatorics*, Cambridge University Press, 2006.

[100] R. Thom, Sur l'homologie des variétés algebriques réelles, in: S.S. Cairns (ed.), *Differential and Combinatorial Topology*, Princeton University Press, Princeton, NJ, 1965, 255–265.

[101] C. D. Tóth, The Szemerédi-Trotter Theorem in the complex plane, *Combinatorica*, **35** (2015), 95–126.

[102] P. Valtr, Strictly convex norms allowing many unit distances and related touching questions, *unpublished manuscript*, 2005.

[103] L. Vinh, Szemerédi-Trotter type theorems and sum-product estimates in finite fields, *European Journal of Combinatorics* **32** (2011), 1177–1181.

[104] H. E. Warren, Lower bound for approximation by nonlinear manifolds, *Trans. Amer. Math. Soc.* **133** (1968), 167–178.

[105] D. B. West, *Introduction to graph theory*, Upper Saddle River: Prentice hall, 2001.

[106] T. Wolff, Recent work connected with the Kakeya problem, *Prospects in mathematics* **2** (1999), 129–162.

[107] J. Zahl, An improved bound on the number of point-surface incidences in three dimensions, *Contrib. Discrete Math.*, **8** (2013), 100–121.

[108] J. Zahl, A Szemeredi-Trotter type theorem in $\mathbb{R}^4$, *Discrete. Comput. Geom.* **54** (2015), 513–572.

[109] J. Zahl, A note on rich lines in truly high dimensional sets, *FoM, Sigma* **4** (2016), 1–13.

[110] F. de Zeeuw, A short proof of Rudnev's point-plane incidence bound, arXiv:1612.02719.

[111] R. Zippel, Probabilistic algorithms for sparse polynomials, *Proc. of the International Symposiumon on Symbolic and Algebraic Computation* (1979), 216–226.