UNIVERSITY LECTURE SERIES VOLUME 64

Polynomial Methods in Combinatorics

Larry Guth

American Mathematical Society

https://doi.org/10.1090//ulect/064

Polynomial Methods in Combinatorics

UNIVERSITY LECTURE SERIES VOLUME 64

Polynomial Methods in Combinatorics

Larry Guth

American Mathematical Society
Providence, Rhode Island

2010 *Mathematics Subject Classification.* Primary 05D99.

For additional information and updates on this book, visit
www.ams.org/bookpages/ulect-64

Library of Congress Cataloging-in-Publication Data

Names: Guth, Larry, 1977–
Title: Polynomial methods in combinatorics / Larry Guth.
Description: Providence, Rhode Island : American Mathematical Society, [2016] | Series: University lecture series ; volume 64 | Includes bibliographical references.
Identifiers: LCCN 2016007729 | ISBN 9781470428907 (alk. paper)
Subjects: LCSH: Combinatorial geometry. | Polynomials. | Geometry, Algebraic. | AMS: Combinatorics – Extremal combinatorics – None of the above, but in this section. msc
Classification: LCC QA167.G88 2016 | DDC 511/.66–dc23 LC record available at http://lccn.loc.gov/2016007729

---

10 9 8 7 6 5 4 3 2 1 21 20 19 18 17 16

Contents

Preface

This book explains some recent progress in combinatorial geometry that comes from an unexpected connection with polynomials and algebraic geometry. One of the early results in this story is a two-page solution of a problem called the finite field Kakeya problem, which experts had believed was extremely deep. The most well-known result in this book is an essentially sharp estimate for the distinct distance problem in the plane, a famous problem raised by Paul Erdős in the 1940s. The book also emphasizes connections between different fields of mathematics. For example, some of the new proofs in combinatorics that we study were suggested by ideas from error-correcting codes. We discuss this connection, as well as related ideas in Fourier analysis, number theory, and differential geometry. First- or second-year graduate students, as well as advanced undergraduates and researchers, should find this book accessible.

My own work in this area is mostly joint with Nets Katz, and I learned a lot about this circle of ideas talking with him and exploring together. I taught a class on this material at MIT in the fall of 2012. I want to thank the students in the class who typed up notes for some of the lectures. Those lecture notes formed a first draft for the book. The students were Sam Elder, Andrey Grinshpun, Nate Harmon, Adam Hesterberg, Chiheon Kim, Gaku Liu, Laszlo Lovasz, Rik Sengupta, Efrat Shaposhnik, Sean Simmons, Yi Sun, Adrian Vladu, Ben Yang, and Yufei Zhao. I also want to thank the following people for looking at drafts of the book and making helpful suggestions: Josh Zahl, Thao Do, Hong Wang, Ben Yang, and Jiri Matoušek. While I was writing the book, I was supported by a Sloan fellowship and a Simons Investigator award.

Finally, I would like to thank my family for their love and support.

Larry Guth, MIT

https://doi.org/10.1090//ulect/064/01

CHAPTER 1

Introduction

This book is about some applications of polynomials to problems in combinatorics. What I think is interesting about these arguments is that the statements of the problems do not involve polynomials, but polynomials provide a crucial structure under the surface. The starting point of the book is Dvir's solution of the finite field Kakeya problem [**D**]. This is a problem on the border between combinatorics and harmonic analysis. People in the field had believed that it was a very hard problem, but the proof is only a few pages long, and it only requires an undergraduate background to understand.

Here is the statement of the finite field Kakeya problem. Let $\mathbb{F}_q$ denote the finite field with q elements. A set $K \subset \mathbb{F}_q^n$ is called a Kakeya set if it contains a line in every direction. (In other words, K is a Kakeya set if it contains a translate of every 1-dimensional subspace of $\mathbb{F}_q^n$.) The question is, what is the smallest possible cardinality of a Kakeya set $K \subset \mathbb{F}_q^n$?

THEOREM 1.1. ([**D**]) If $K \subset \mathbb{F}_q^n$ is a Kakeya set, then

$$|K| \geq (10n)^{-n} q^n.$$

For a fixed dimension n, this estimate says that the size of a Kakeya set is at least a constant factor times the size of the whole vector space $\mathbb{F}_q^n$.

The proof begins by considering a lowest degree (non-zero) polynomial that vanishes on the set K. Using this polynomial, the proof is short and clean. But without mentioning this polynomial, proving Theorem 1.1 seems to be very difficult, and people worked hard to prove much weaker estimates. Over the course of the book, we will explore different variations of this trick with polynomials, and we will discuss applications of this method to different problems.

A large piece of the book is about the distinct distance problem in the plane, a combinatorics problem raised by Erdős in the 1940s [**Er1**]. The problem asks, given a set of N points in the plane, what is the minimum possible number of distinct distances between the points. For example, if the points are evenly spaced along a line, then there are $N - 1$ distinct distances. Erdős checked that arranging the points in a square grid is slightly more efficient, giving on the order of $N(\log N)^{-1/2}$ distinct distances. He conjectured that grids are nearly optimal. We will prove the following lower bound, which nearly matches the grid example:

THEOREM 1.2. (Guth and Katz, [**GK2**], building on [**ElSh**]) If P is a set of N points in $\mathbb{R}^2$, then the number of distinct distances between the points of P is at least $cN(\log N)^{-1}$.

The main applications in the book are to problems in combinatorics. But it is also striking to me that this trick with polynomials has connections to several

other areas of mathematics. We will see related arguments connected with error-correcting codes in computer science, inequalities about surface area in differential geometry, diophantine equations in number theory, and geometric estimates related to Fourier analysis. Each of these arguments has some significant ingredients in common with the proof of the finite field Kakeya conjecture. Also, each of these fields offers its own perspective about why polynomials are special functions and what makes them useful in these applications.

I tried to make the book self-contained, and I hope that it will be accessible to readers with a first-year graduate background or a strong undergraduate background.

In the rest of the introduction, we give an overview of the book. Some readers might want to begin by reading the overview. Other readers might want to begin by reading the proof of finite field Kakeya in Chapter 2.

1.1. Incidence geometry

When the distinct distance problem was first raised, in [**Er1**] in the 1940's, it didn't fit into any well-developed field of mathematics. There were a small number of isolated problems of this flavor that different people had raised in different situations. Some of these problems - including the distinct distance problem - turned out to be surprisingly hard. Over the next few decades, people began to study this circle of problems in a systematic way, and they developed a field of combinatorics called incidence geometry. Broadly speaking, incidence geometry is the study of combinatorial problems about basic geometric objects, like lines, circles, angles, or distances. To give a little sense of this area, let us describe one of the important theorems and some open problems.

One fundamental question in the field has to do with the possible intersection patterns of lines in the plane. If $\mathfrak{L}$ is a set of lines, a point x is called an r-rich point of $\mathfrak{L}$ if x lies in at least r lines of $\mathfrak{L}$. The set of r-rich points of $\mathfrak{L}$ is denoted $P_r(\mathfrak{L})$. Given a certain number of lines, how many r-rich points can they form? In the early 1980s, Szemerédi and Trotter solved this problem up to a constant factor.

THEOREM 1.3. ([**SzTr**], 1983) There are constants $0 < c < C$ so that the following holds. If $2 \leq r \leq L^{1/2}$, then

$$cL^2 r^{-3} \leq \max_{|\mathfrak{L}|=L} |P_r(\mathfrak{L})| \leq CL^2 r^{-3}.$$

If $L^{1/2} \leq r \leq L$, then

$$cLr^{-1} \leq \max_{|\mathfrak{L}|=L} |P_r(\mathfrak{L})| \leq CLr^{-1}.$$

The lower bound comes from a fairly simple example involving a grid of points. The difficult part is the upper bound. A remarkable thing about this proof is that it is based on topology. The topological approach was developed further by other mathematicians in the field, in papers such as [**CEGSW**] and [**Sz**], leading to a range of tools that apply to many problems. Developing topological methods to prove combinatorial estimates of this kind is one of the main achievements of incidence geometry.

Incidence geometry also has many simply stated open problems. For instance, in Theorem 1.3, if we replace lines by circles, we get a difficult open problem. Replacing lines by unit circles gives a different difficult open problem. Replacing

lines by ellipses or parabolas gives two more difficult open problems. These problems have been studied intensively for decades. The topological methods discussed above give interesting bounds for these problems, but the best current bounds don't match any known examples, and most people believe the bounds are not sharp. The distinct distance problem was also studied by these topological methods. People proved interesting bounds, but the method runs into similar issues as it does in problems about circles.

In the last decade, polynomial methods have developed into a second major approach to incidence geometry. Here is an example of an incidence geometry problem that seemed out of reach with topological methods but which has a short proof using polynomials. The joints problem is a problem about lines in $\mathbb{R}^3$, which was raised in the early 90s by [**CEGPSSS**]. If $\mathfrak{L}$ is a set of lines in $\mathbb{R}^3$, then a point x is a joint of $\mathfrak{L}$ if x lies in three non-coplanar lines of $\mathfrak{L}$. It is not hard to find examples with L lines and on the order of $L^{3/2}$ joints, and [**CEGPSSS**] conjectured that the number of joints is always at most $CL^{3/2}$. Before the polynomial method, the best known bound was $L^{1.62\ldots}$ ([**FS**]) and the argument was fairly complex.

THEOREM 1.4. ([**GK1**], proof simplified by [**KSS**], [**Q**]) A set of L lines in $\mathbb{R}^3$ forms at most $CL^{3/2}$ joints.

We will prove this result in Chapter 2, right after the proof of the finite field Kakeya conjecture.

In [**ElSh**], Elekes and Sharir proposed a new approach to the distinct distance problem, which connects it to the incidence geometry of lines in $\mathbb{R}^3$. This approach led to new questions about lines in $\mathbb{R}^3$, which I think are natural questions in their own right. These questions were resolved by Nets Katz and the author in [**GK2**]. The proofs use polynomial methods, and they also bring into play the topological methods described above and more tools from algebraic geometry. Explaining these results and their applications is one of the main goals of the book.

Suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$. Let us first consider the 2-rich points of $\mathfrak{L}$. Since any two lines intersect in at most one point, the number of 2-rich points is at most $\binom{L}{2}$, and this can actually happen if all the lines lie in a plane. But the scenario that all lines lie in a plane is a special situation. If we rule out this situation, can we get a better bound? For instance, in the approach to the distinct distance problem from [**ElSh**], one is led to a set $\mathfrak{L}$ of L lines in $\mathbb{R}^3$ with at most $L^{1/2}$ lines in any plane. For such a set, can we prove a significantly stronger bound for $|P_2(\mathfrak{L})|$?

Suprisingly the answer is no. The counterexample comes from a degree 2 algebraic surface, such as the surface defined by $z = xy$. This surface is doubly ruled – every point in the surface lies in two lines in the surface. Choosing L lines contained in this degree 2 surface, we get a set $\mathfrak{L}$ with $L^2/4$ 2-rich points, while any plane contains at most 2 lines of $\mathfrak{L}$. This doubly ruled surface has been known in algebraic geometry for a long time, and it plays an important role in the first paper on the joints problem [**CEGPSSS**]. This example, involving a polynomial surface, helps to explain why polynomials play an important role in studying the intersection patterns of lines in space.

What if we assume that $\mathfrak{L}$ contains at most $L^{1/2}$ lines in any plane or any degree 2 algebraic surface? With these stronger hypotheses, can we prove a significantly stronger bound on $|P_2(\mathfrak{L})|$? This time, the answer is yes. The methods from [**CEGPSSS**] give a significant improvement, and [**GK2**] gives a sharp estimate.

THEOREM 1.5. If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$, and at most $L^{1/2}$ lines of $\mathfrak{L}$ lie in any plane or any degree 2 algebraic surface, then

$$|P_2(\mathfrak{L})| \leq CL^{3/2}.$$

The proof of Theorem 1.5 uses polynomial methods, and it also draws on the theory of ruled surfaces from algebraic geometry. (An algebraic surface is called ruled if it contains a line through every point.)

What about r-rich points for $r > 2$? If all the lines of $\mathfrak{L}$ lie in a plane, then the Szemerédi-Trotter theorem gives a sharp upper bound for $|P_r(\mathfrak{L})|$. We focus on the range $3 \leq r \leq L^{1/2}$, which is more challenging and interesting. In this range, Theorem 1.3 gives

$$|P_r(\mathfrak{L})| \leq CL^2 r^{-3}.$$

It's not hard to extend this bound to any set of L lines in $\mathbb{R}^3$. But suppose that $\mathfrak{L}$ contains at most $L^{1/2}$ lines in any plane. Can we prove a significantly better upper bound? The answer is yes, and the following sharp upper bound was proven in [**GK2**].

THEOREM 1.6. ([**GK2**]) If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$, and at most $L^{1/2}$ lines of $\mathfrak{L}$ lie in any plane, and if $3 \leq r \leq L^{1/2}$ then

$$|P_r(\mathfrak{L})| \leq CL^{3/2} r^{-2}.$$

The proof of Theorem 1.6 combines polynomial methods with topological methods that come from the proof of Theorem 1.3.

Theorems 1.5 and 1.6 give a lot of understanding of the incidence geometry of lines in $\mathbb{R}^3$. The distinct distance estimate, Theorem 1.2, follows by combining them with the framework from [**ElSh**].

1.2. Connections with other areas

The proofs of these combinatorial results have some similarities to proofs in other fields, and we will discuss a number of these connections.

One connection involves error-correcting codes in computer science. Dvir's background is in computer science. His interests include error-correcting codes, and perspectives from coding theory helped lead to the proof of finite field Kakeya. Here is a typical problem from error-correcting codes. Suppose that $\mathbb{F}_q$ is the finite field with q elements and that $Q : \mathbb{F}_q \to \mathbb{F}_q$ is a polynomial whose degree is not too high. Suppose that we have access to a corrupted version of Q. More precisely, suppose that $F : \mathbb{F}_q \to \mathbb{F}_q$ is a function which is known to agree with Q for a certain fraction of points $x \in \mathbb{F}_q$. By looking at F, we would like to recover the original polynomial Q, and we would like to do so efficiently. Berlekamp and Welch [**BW**] discovered an interesting trick for recovering the original polynomial, and this trick has common ingredients with the proof of finite field Kakeya.

It turns out that even if F contains quite a lot of corruption, it is still possible to efficiently recover the polynomial Q. In the field of error-correcting codes, polynomials are known for their resiliency - a polynomial code can tolerate a high level of error, and the original information can still be recovered. Polynomials are important in error-correcting codes because they are an especially resilient class of functions in this sense.

In differential geometry, polynomials are known for their efficiency. There are many examples of this efficiency. To mention one classical example, the zero set

of a complex polynomial is an area-minimizing surface – a surface with the least possible area given its boundary. Most of these results about efficiency involve very different ideas from the ones in this book, but there is one recent example involving closely related ideas. This result is a theorem of Gromov [**Gr**] proving surface area estimates for families of functions. It takes a little work to set up the statement of this theorem, so we postpone it to the chapter on polynomial methods in differential geometry. Roughly speaking, the theorem says that polynomials are a particularly efficient class of functions in terms of the surface areas of their zero sets. The proof from [**Gr**] has a parallel structure to the proof of finite field Kakeya. It also involves a different tool coming from topology, the polynomial ham sandwich theorem. This tool coming from the geometry literature plays a role in the proof of the distinct distance estimate.

To summarize the last few paragraphs, polynomials are efficient from the point of view of differential geometry, and polynomials are resilient from the point of view of error-correcting codes. These two facts are related to each other, and the proofs in both fields share some common ingredients with the proof of finite field Kakeya.

A third connection involves diophantine equations in number theory. In the early 20th century, Thue proved that a broad class of diophantine equations in two variables have only finitely many integer solutions. His theorem was important in part because it covers a much broader class of equations than any previous work in the subject. Here is the statement of the theorem.

THEOREM 1.7. Suppose that $P(x, y) \in \mathbb{Z}[x, y]$ is a homogeneous polynomial of degree $d \geq 3$ which is irreducible over $\mathbb{Z}$. (For instance $P(x, y) = y^d - 2x^d$ for $d \geq 3$.) Then, for any integer A, the diophantine equation $P(x, y) = A$ has only finitely many integer solutions $(x, y) \in \mathbb{Z}^2$.

The proof of Theorem 1.7 also involves some similar ideas to the proof of finite field Kakeya. The statement of Theorem 1.7 involves a polynomial $P(x, y)$, but the proof also involves a lot of other polynomials, called auxiliary polynomials. The auxiliary polynomials in the proof play a similar role to the polynomial in the proof of finite field Kakeya. In the chapter on diophantine equations, we will prove Theorem 1.7 and discuss the parallels with the other proofs in the book.

Finally, we mention the original Kakeya problem. The finite field Kakeya problem was invented as a cousin for the original Kakeya problem, which involves the behavior of lines in $\mathbb{R}^n$. Recall that a finite field Kakeya set $K \subset \mathbb{F}_q^n$ is a set which contains a line in every direction. Similarly, a Kakeya set $K \subset \mathbb{R}^n$ is a set which contains a unit line segment in every direction. There are several variations of the Kakeya problem, but they all have to do with how big a Kakeya set needs to be. For instance, one version asks about the minimum possible Hausdorff dimension of a Kakeya set. All known Kakeya sets $K \subset \mathbb{R}^n$ have Hausdorff dimension n. The Kakekya problem is about the possible intersection patterns of lines in $\mathbb{R}^n$, but unlike in incidence geometry, we consider infinitely many lines instead of finitely many lines. The Kakeya problem can also be rephrased in terms of the intersection patterns of finitely many thin tubes in $\mathbb{R}^n$. This description in terms of thin tubes is the most useful for working on the problem and also the most useful in applications, so we emphasize it in the book. Here is a version of the Kakeya problem involving the intersection patterns of long thin tubes.

QUESTION 1.8. Fix a dimension n and let $\delta > 0$ be a small parameter. Suppose that $\mathcal{T}$ is a set of cylindrical tubes in $\mathbb{R}^n$, each of radius $\delta > 0$ and length 1. For a tube T, let $v(T)$ be a unit vector in the direction of T. Suppose that for any two tubes $T_1, T_2 \in \mathcal{T}$, the angle between $v(T_1)$ and $v(T_2)$ is at least δ, and suppose that for any unit vector w, there is some $T \in \mathcal{T}$ so that the angle between $v(T)$ and w is at most 10δ. What is the minimum possible volume of the union of the tubes of $\mathcal{T}$?

If the tubes of $\mathcal{T}$ are disjoint, then it is easy to check that the volume of the union is on the order of 1. In the 1920s, Besicovitch constructed an ingenious example where the volume of the union goes to zero with δ. A slightly improved version of this construction [**Sch**] gives logarithmic decay:

$$|\cup_{T\in\mathcal{T}} T| \leq C_n \frac{1}{|\log \delta|}.$$

This construction is still the best one known. The Kakeya conjecture asserts that, for any $\varepsilon > 0$, the volume of the union of the tubes in $\mathcal{T}$ is at least $c(\varepsilon)\delta^{\varepsilon}$. The best known lower bounds for the volume are much weaker: for instance, if $n = 3$, we know that the volume of the union is at least $c\delta^{1/2}$.

In the 1970's, mathematicians discovered that the Kakeya problem is closely connected to a circle of problems in Fourier analysis. This connection encouraged a lot of interest in the problem, and it has been studied intensively ever since.

It is not clear how much the polynomial method can contribute to the original Kakeya problem. The proof of finite field Kakeya seems like it might be an important clue, but there are major difficulties in trying to adapt the proof from lines in $\mathbb{F}_q^n$ to thin tubes in $\mathbb{R}^n$. On the other hand, the polynomial method has had some successes proving harmonic analysis estimates related to the Kakeya problem. We will discuss all these issues in the chapter on harmonic analysis.

1.3. Outline of the book

The first part of the book is about introducing the polynomial methods we will study. In Chapter 2, we prove the finite field Kakeya theorem and the joints theorem, and we outline the ingredients of the method. In Chapter 3, we discuss why these problems were difficult to solve without polynomials and what features of polynomials make them useful. The proofs in Chapter 2 are partly based on ideas from error-correcting codes. In Chapter 4, we study the Berlekamp-Welch algorithm and other work in error-correcting codes, and we see how it relates to the proofs in Chapter 2. In Chapter 5, we discuss some earlier work in combinatorics with a similar flavor. In Chapter 6, we prove the Bezout theorem, a fundamental theorem of algebraic geometry. We will use this result in the later chapters, and we also give a proof with a somewhat similar flavor to the other proofs in the book.

The second part of the book gives background in incidence geometry. In Chapter 7, we prove the Szemerédi-Trotter theorem, and introduce some of the topological methods in the area. We discuss the distinct distance problem as well as some hard open problems in the field. In Chapter 8, we discuss incidence geometry in higher dimensions, especially dimension three. In Chapter 9, we discuss the partial symmetry approach to the distinct distance problem.

The third part of the book is about applications of the polynomial method in incidence geometry. In this part of the book, we prove Theorems 1.5 and 1.6. These

proofs involve several different tools, and we introduce one tool in each chapter. Chapter 10 introduces polynomial partitioning. This is an important tool, and it turns out to be enough to prove a slightly weaker form of the distinct distance estimate. Chapter 11 explores the connection between combinatorial structure and algebraic structure. Chapter 12 combines these tools and finishes the proof of Theorem 1.6. Chapter 13 introduces tools from ruled surface theory in algebraic geometry and proves Theorem 1.5.

The fourth part of the book discusses connections with a few other areas. Chapter 14 discusses connections with differential geometry. Chapter 15 discusses the Kakeya problem and Fourier analysis. Chapter 16 discusses Thue's work on diophantine equations.

1.4. Other connections between polynomials and combinatorics

There are a lot of interesting connections between polynomials and combinatorics. I wanted to mention a few interesting directions that have a similar flavor to the topics in this book.

The book *Linear algebra methods in combinatorics* [**BF**], by Babai and Frankl, develops a circle of ideas involving polynomials, linear algebra, and combinatorics. The recent book by Matousek, [**Ma2**], discusses many of the same ideas. We will touch on this work briefly in Chapter 5.

Alon proved a variant of the Hilbert Nullstellensatz from algebraic geometry, called the combinatorial nullstellensatz. This is a theorem about polynomials, related to classical theorems of Chevalley and Warning. He and others applied this theorem to problems in combinatorics, including some additive number theory and some graph theory. See the survey [**Al**] and the references therein for an introduction to this area.

Recently, Green and Tao [**GT**] proved some old conjectures in incidence geometry with an argument that uses a combination of topology and polynomials. We will say more about these results in Section 7.5.

1.5. Notation

I would like to introduce one convenient piece of notation here. We write $A \lesssim B$ to mean that there is some constant C so that $A \leq CB$. We write $A \sim B$ to mean that $A \lesssim B$ and $B \lesssim A$.

We will introduce other notation as it comes up, but we record here for reference a few basic pieces of notation that will come up a lot in the book. We let $\mathbb{F}$ denote a field, and we let $\mathbb{F}_q$ denote the finite field with q elements. We let $\mathrm{Poly}_D(\mathbb{F}^n)$ be the space of polynomials in n variables, with coefficients in $\mathbb{F}$, and with total degree at most D. If P is a polynomial, then we write $Z(P)$ for the zero set of P. If $\mathfrak{L}$ is a set of lines, then we write $P_r(\mathfrak{L})$ for the set of r-rich points of $\mathfrak{L}$ - the set of points that lie in at least r lines of $\mathfrak{L}$.

https://doi.org/10.1090//ulect/064/02

CHAPTER 2

Fundamental examples of the polynomial method

In this chapter, we give our first examples of the polynomial method, proving the finite field Kakeya theorem and the joints theorem. We begin by introducing two simple tools that we will use all through the book, called the parameter counting argument and the vanishing lemma. These tools are simple lemmas from undergraduate algebra. Using these tools, we can prove the theorems above in about one page each.

2.1. Parameter counting arguments

Let $\mathbb{F}$ be a field. Let $\mathrm{Poly}_D(\mathbb{F}^n)$ be the space of polynomials in n variables, with coefficients in $\mathbb{F}$, and with total degree at most D. If the n variables are $x_1, ..., x_n$, then $\mathrm{Poly}_D(\mathbb{F}^n)$ is the subset of the polynomial ring $\mathbb{F}[x_1, ..., x_n]$ consisting of polynomials of degree at most D. The space $\mathrm{Poly}_D(\mathbb{F}^n)$ is a vector space over the field $\mathbb{F}$.

Suppose that $S \subset \mathbb{F}^n$ is a finite set. We would like to know if there is a non-zero polynomial $P \in \mathrm{Poly}_D(\mathbb{F}^n)$ that vanishes on the set S. We can get some basic information on this question by a dimensional argument.

PROPOSITION 2.1. If $\mathrm{Dim}\, \mathrm{Poly}_D(\mathbb{F}^n) > |S|$, then there is a non-zero polynomial $P \in \mathrm{Poly}_D(\mathbb{F}^n)$ which vanishes on S.

PROOF. Let $p_1, ..., p_{|S|}$ be the points of S. We let E be the evaluation map $E : \mathrm{Poly}_D(\mathbb{F}^n) \to \mathbb{F}^{|S|}$ defined by

$$E(Q) = \big(Q(p_1), ..., Q(p_{|S|})\big)\,.$$

The map E is a linear map. (It follows from the formula for E that for any $Q_1, Q_2 \in \mathrm{Poly}_D(\mathbb{F}^n)$, $E(Q_1 + Q_2) = E(Q_1) + E(Q_2)$ and for any $Q \in \mathrm{Poly}_D(\mathbb{F}^n), \lambda \in \mathbb{F}$, $E(\lambda Q) = \lambda E(Q)$.) The kernel of the map E is exactly the set of polynomials in $\mathrm{Poly}_D(\mathbb{F}^n)$ that vanish on S.

If $\mathrm{Dim}\, \mathrm{Poly}_D(\mathbb{F}^n) > |S|$, then the dimension of the domain of E is greater than the dimension of the target of E. By the rank-nullity theorem from linear algebra, the map E must have a non-trivial kernel. Therefore, there is a non-zero polynomial $P \in \mathrm{Poly}_D(\mathbb{F}^n)$ that vanishes on S. □

This result raises the question "what is the dimension of $\mathrm{Poly}_D(\mathbb{F}^n)$?" A basis for $\mathrm{Poly}_D(\mathbb{F}^n)$ is given by the monomials $x_1^{D_1}...x_n^{D_n}$ where D_i are non-negative integers with $\sum D_i \leq D$. By counting the elements of this basis, we can compute the dimension of $\mathrm{Poly}_D(\mathbb{F}^n)$.

LEMMA 2.2. The dimension of $\mathrm{Poly}_D(\mathbb{F}^n)$ is $\binom{D+n}{n}$. In particular,

$$\mathrm{Dim}\, \mathrm{Poly}_D(\mathbb{F}^n) \geq D^n/n!$$

Heuristically, we have n exponents D_i to choose. Each exponent has $\sim D$ choices, and so we expect $\operatorname{Dim} \operatorname{Poly}_D(\mathbb{F}^n) \sim D^n$. For example, we can choose D_i to be any integer in the range $0 \leq D_i \leq D/n$, and so $\operatorname{Dim} \operatorname{Poly}_D(\mathbb{F}^n) \geq D^n n^{-n}$. This crude estimate is strong enough for our applications, but we also explain how to compute the dimension precisely.

PROOF. Fix D and n. We encode a monomial $x_1^{D_1} ... x_n^{D_n}$ by a string of D $*$'s and n $|$'s as follows. We begin with D_1 $*$'s. Then we put one $|$. Then we put D_2 $*$'s. Then we put a second $|$. We continue this way until we have put D_n $*$'s, followed by an n^{th} $|$. Finally we put $D - \sum_i D_i$ $*$'s. This encoding is a bijection between all the monomials in $\operatorname{Poly}_D(\mathbb{F}^n)$ and all the strings of D $*$'s and n $|$'s. Therefore, the number of monomials is $\binom{D+n}{n}$. □

Plugging in this information about $\operatorname{Dim} \operatorname{Poly}_D(\mathbb{F}^n)$ into Proposition 2.1, we get the following result.

LEMMA 2.3. (Parameter counting) If $S \subset \mathbb{F}^n$ and $|S| < \binom{D+n}{n}$, then there is a non-zero polynomial $P \in \operatorname{Poly}_D(\mathbb{F}^n)$ that vanishes on S.

The following corollary is a little less sharp, but it's a useful way to encapsulate the information.

LEMMA 2.4. For any $n \geq 2$, for any finite set $S \subset \mathbb{F}^n$, there is a non-zero polynomial that vanishes on S with degree $\leq n|S|^{1/n}$.

PROOF. Define D to be the greatest integer $\leq n|S|^{1/n}$. By an elementary calculation $\binom{D+n}{n} > |S|$. Now by the last Corollary, there is a non-zero $P \in \operatorname{Poly}_D(\mathbb{F}^n)$ that vanishes on S. □

2.2. The vanishing lemma

Here is a fundamental fact about polynomials that will play a role throughout the book.

LEMMA 2.5. If $P \in \operatorname{Poly}_D(\mathbb{F})$, and if P vanishes at $D+1$ points, then P is the zero polynomial.

This lemma should be familiar to most readers, but we include the proof because the lemma plays such an important role in the book. We begin with a couple of other basic lemmas about polynomials which we use to prove Lemma 2.5.

LEMMA 2.6. If $P(x) \in \operatorname{Poly}_D(\mathbb{F})$ is a polynomial in one variable and $x_1 \in \mathbb{F}$, then we can write P in the form

$$P(x) = (x - x_1)P_1(x) + r,$$

where $P_1(x) \in \operatorname{Poly}_{D-1}(\mathbb{F})$ and $r \in \mathbb{F}$.

PROOF. We do the proof by induction on D. If $D = 0$, then P is a constant and the conclusion is clear.

Suppose $P(x) = \sum_{j=0}^{D} a_j x^j$. Let $Q(x) = P(x) - (x - x_1)(a_D x^{D-1})$. The x^D term of $Q(x)$ vanishes, and so $Q(x) \in \operatorname{Poly}_{D-1}(\mathbb{F})$. By induction, we have

$$P(x) - (x - x_1)(a_D x^{D-1}) = Q(x) = (x - x_1)Q_1(x) + r,$$

where $Q_1 \in \operatorname{Poly}_{D-2}(\mathbb{F})$ and $r \in \mathbb{F}$. Therefore, we see

$$P(x) = (x - x_1)(a_D x^{D-1} + Q_1(x)) + r.$$ □

LEMMA 2.7. If $P(x) \in \mathrm{Poly}_D(\mathbb{F})$ is a polynomial over a field $\mathbb{F}$ and $P(x_1) = 0$ for some $x_1 \in \mathbb{F}$, then $P(x) = (x - x_1)P_1(x)$ for some polynomial $P_1 \in \mathrm{Poly}_{D-1}(\mathbb{F})$.

PROOF. By the previous lemma, we can write P in the form $P(x) = (x - x_1)P_1(x) + r$. Plugging in $P(x_1) = 0$, we see that $r = 0$. □

Now we are ready to prove Lemma 2.5:

PROOF. We prove the lemma by induction on D. As a base case suppose that $D = 0$ so that P is a constant. If P vanishes at one point, then P must be the zero polynomial.

Now we do the inductive step. Suppose that $P \in \mathrm{Poly}_D(\mathbb{F})$ and that P vanishes at $D + 1$ distinct points $x_1, ..., x_{D+1}$. By Lemma 2.7, we see that there is some polynomial $P_1 \in \mathrm{Poly}_{D-1}(\mathbb{F})$ so that

$$P(x) = (x - x_{D+1})P_1(x).$$

But P_1 must vanish at $x_1, ..., x_D$. By the inductive hypothesis, it follows that $P_1 = 0$, and so $P = 0$. □

A line $l \subset \mathbb{F}^n$ is a 1-dimensional affine subspace.

LEMMA 2.8. (Vanishing lemma) If $P \in \mathrm{Poly}_D(\mathbb{F}^n)$ and P vanishes at $D + 1$ points on a line l, then P vanishes at every point of l.

PROOF. We parametrize l by a map $\gamma : \mathbb{F} \to \mathbb{F}^n$ of the form $\gamma(t) = at + b$, for vectors $a, b \in \mathbb{F}^n$, with $a \neq 0$. We define $Q(t) = P(\gamma(t)) = P(at + b)$. We see that $Q(t)$ is a polynomial in one variable of degree $\leq D$. Since P vanishes at $D + 1$ points of l, Q vanishes at $D + 1$ values of t. By Lemma 2.5, Q is the zero polynomial, and so P vanishes on l. □

2.3. The finite-field Nikodym problem

Let $\mathbb{F}_q$ be a finite field with q elements. A set $N \subset \mathbb{F}_q^n$ is called a Nikodym set if, for each point $x \in \mathbb{F}_q^n$, there is a line $L(x)$ containing x so that $L(x) \setminus \{x\} \subset N$. A trivial example of a Nikodym set is the entire set $\mathbb{F}_q^n$. Can one find a significantly smaller Nikodym set?

THEOREM 2.9. ([**D**]) Any Nikodym set in $N \subset \mathbb{F}_q^n$ contains at least $c_n q^n$ elements. We can take $c_n = (10n)^{-n}$.

PROOF. We do a proof by contradiction. Let us assume that N is a Nikodym set with $|N| < (10n)^{-n} q^n$. By the parameter counting argument, Lemma 2.4, we can find a non-zero polynomial P that vanishes on N with degree bounded by

$$\mathrm{Deg}\, P \leq 2n|N|^{1/n} \leq (1/5)q < q - 1.$$

Next we claim that P vanishes at every point of $\mathbb{F}_q^n$. Let x be an arbitrary point of $\mathbb{F}_q^n$. By the definition of a Nikodym set, there is a line $L(x)$ containing x so that $L(x) \setminus \{x\} \subset N$. The polynomial P vanishes on N, so P vanishes at $\geq q - 1$ points of $L(x)$. Since $\mathrm{Deg}\, P < q - 1$, the vanishing lemma implies that P vanishes on the whole line $L(x)$. In particular, P vanishes at x.

We know that P is a non-zero polynomial and that P vanishes at every point of $\mathbb{F}_q^n$. That might sound like a contradiction, but it's not quite a contradiction. For example, the polynomial $x^q - x$ vanishes for all $x \in \mathbb{F}_q$. But we also know that $\mathrm{Deg}\, P < q$. We now get a contradiction from the following lemma.

LEMMA 2.10. Suppose that $P \in \mathrm{Poly}_{q-1}(\mathbb{F}_q^n)$. If P vanishes at every point of $\mathbb{F}_q^n$, then P is the zero polynomial.

PROOF. The proof uses the vanishing lemma and induction on n.

If $n = 1$, then the result follows directly from the vanishing lemma, Lemma 2.5. P vanishes at $q > \mathrm{Deg}\, P$ points of $\mathbb{F}_q$, and so P is the zero polynomial.

Now we proceed by induction. We let $x_1, ..., x_n$ be coordinates on $\mathbb{F}_q^n$, and we write P in the form

$$P(x_1, ..., x_n) = \sum_{j=0}^{q-1} P_j(x_1, ..., x_{n-1}) x_n^j.$$

In this formula, P_j are polynomials in $x_1, ..., x_{n-1}$ of degree $\leq q-1$. Fix the values of $x_1, ..., x_{n-1}$, and let x_n vary. We have a polynomial in x_n, of degree $\leq q-1$, that vanishes for all $x_n \in \mathbb{F}_q$. By Lemma 2.5, this polynomial must be the zero polynomial. In other words, $P_j(x_1, ..., x_{n-1}) = 0$ for all j and all $(x_1, ..., x_{n-1}) \in \mathbb{F}_q^{n-1}$. But now, by induction on n, each polynomial P_j is the zero polynomial. Then P is the zero polynomial as well. □

This finishes the proof of the finite field Nikodym theorem. □

Here is a summary of the proof. Suppose that N is a small Nikodym set. By parameter counting, we can find a polynomial that vanishes on N with degree less than $q-1$. By the vanishing lemma, this polynomial vanishes at every point of $\mathbb{F}_q^n$. Now the polynomial vanishes too much, giving a contradiction.

Before this approach with polynomials, the finite field Nikodym problem was considered to be very difficult. The finite field Nikodym problem has a more well-studied cousin called the finite field Kakeya problem. The paper [**D**] solved this problem as well with a small variation on the argument above. We give the proof in the next section.

2.4. The finite field Kakeya problem

A set $K \subset \mathbb{F}_q^n$ is called a Kakeya set if it contains a line in every direction. In other words, for every vector $a \in F^n \setminus \{0\}$, there is a vector b so that the line $\{at + b | t \in \mathbb{F}_q\}$ is contained in K. A trivial example of a Kakeya set is the entire vector space $\mathbb{F}_q^n$. Can one find a Kakeya set significantly smaller than this?

THEOREM 2.11. ([**D**]) A Kakeya set $K \subset \mathbb{F}_q^n$ has at least $c_n q^n$ elements, for $c_n = (10n)^{-n}$.

PROOF. Suppose that $K \subset \mathbb{F}_q^n$ is a Kakeya set with $|K| < (10n)^{-n} q^n$. By the parameter counting argument, Lemma 2.4, there is a non-zero polynomial P that vanishes on K with $\mathrm{Deg}\, P \leq n|K|^{1/n} < q$. We write P as a sum of two pieces - the terms of highest degree plus the terms of lower degree. If D is the degree of P, then we have $P = P_D + Q$, where P_D is homogeneous of degree D and $\mathrm{Deg}\, Q < D$, and where P_D is non-zero.

Let a be any non-zero vector in $\mathbb{F}_q^n$. Choose b so that the line $\{at + b | t \in \mathbb{F}\}$ is contained in K. Consider the polynomial in one variable $R(t) := P(at + b)$. The polynomial R vanishes for each $t \in F$. It has degree $\leq D < q$. By the vanishing lemma, Lemma 2.5, R is the zero polynomial. In other words, every coefficient of R is zero. But the coefficient of t^D in R is exactly $P_D(a)$. So we see that $P_D(a)$

vanishes for all $a \in F^n \setminus \{0\}$! Since P_D is homogeneous of degree $D \geq 1$, P_D also vanishes at 0, and so P_D vanishes at every point of $\mathbb{F}_q^n$. Since $D < q$, Lemma 2.10 implies that P_D is the zero polynomial. This gives a contradiction. □

Splitting off the homogeneous part in this way has a geometric interpretation. We recall that the projective space $\mathbb{PF}^n$ is the set of equivalence classes of $\mathbb{F}^{n+1} \setminus \{0\}$ where two elements are equivalent if one is a rescaling of the other by a factor $\lambda \in \mathbb{F}^*$. The projective space $\mathbb{PF}^n$ can be written as a disjoint union $\mathbb{F}^n \cup \mathbb{PF}^{n-1}$. The usual way to do this is to identify a point $(x_1, ..., x_n) \in \mathbb{F}^n$ with the equivalence class of $(x_1, ..., x_n, 1)$ in $\mathbb{PF}^n$. The remainder of $\mathbb{PF}^n$ is the equivalence classes of points of the form $(x_1, ..., x_n, 0)$, and this naturally identifies with $\mathbb{PF}^{n-1}$. The subset $\mathbb{PF}^{n-1} \subset \mathbb{PF}^n$ is called the set of points at infinity.

Every line in $\mathbb{F}^n$ can be extended to a projective line in $\mathbb{PF}^n$ by adding one point at infinity. If $a \neq 0$, then the line $\{at + b | t \in \mathbb{F}\}$ in $\mathbb{F}^n$ extends to include the point at infinity in the equivalence class $(a, 0)$. Similarly, if $P \in \mathrm{Poly}_D(\mathbb{F}^n)$, then the zero set of P, $Z(P) \subset \mathbb{F}^n$, can be naturally extended to $\mathbb{PF}^n$ as follows: if $0 \neq a \in \mathbb{F}_q^n$, then the point at infinity $(a, 0)$ lies in $Z(P)$ if and only if $P_D(a) = 0$. In the proof of finite field Kakeya, we showed that if a line $l \subset \mathbb{F}^n$ lies in the zero set of a polynoimal P of degree $< q$, then the point of l at infinity also lies in $Z(P)$. We can think of this as a version of the vanishing lemma in projective space.

This language is nice for summarizing the proof of Theorem 2.11. Suppose that $K \subset \mathbb{F}_q^n$ is a small Kakeya set. By parameter counting, there is a polynomial that vanishes on K with degree less than q. Since K is a Kakeya set, the polynomial vanishes on one line in every direction. By a version of the vanishing lemma, the polynomial vanishes at all the points at infinity in $\mathbb{PF}_q^n$. But then the polynomial vanishes at too many points at infinity, giving a contradiction.

The Kakeya and Nikodym problems presented here are finite field analogues of deep open problems in Euclidean space $\mathbb{R}^n$. Here is the original Kakeya problem in $\mathbb{R}^n$. A Kakeya set $K \subset \mathbb{R}^n$ is a set which contains a unit line segment in each direction. For example, the ball of radius 1/2 is a Kakeya set. Besicovitch constructed surprising examples of Kakeya sets with arbitrarily small volume and even with measure zero. The sets coming from this construction have measure zero, but they all have full Hausdorff dimension. The Kakeya conjecture states that every Kakeya set $K \subset \mathbb{R}^n$ has Hausdorff dimension n. The Kakeya problem is connected with some other deep problems in Fourier analysis, and it was been studied intensively.

The finite field Kakeya problem was first raised as a cousin of the Euclidean problem. Before the polynomial method, it was generally thought that finite field Kakeya was of the same order of difficulty as the original problem. When this proof appeared, there was a sense of shock in the harmonic analysis community. It remains unclear how much the polynomial approach can contribute to understanding the original problem in $\mathbb{R}^n$. We will return to the original Kakeya problem in Chapter 15 and discuss these issues more.

2.5. The joints problem

Let $\mathfrak{L}$ be a set of lines in $\mathbb{R}^3$. A joint of $\mathfrak{L}$ is a point which lies in three non-coplanar lines of $\mathfrak{L}$. The joints problem asks what is the maximal number of joints that can be formed from L lines.

The joints problem was posed in the early 90's by Chazelle, Edelsbrunner, Guibas, Pollack, Seidel, Sharir, and Snoeyink, in [**CEGPSSS**]. They proved that the number of joints formed by L lines is $\lesssim L^{7/4}$, and the exponent has gradually improved. We will explain some of the ideas from that first paper later in the book, in Section 8.4.

Let us look at some examples. Our first example is based on a grid. Consider an $S \times S \times S$ grid of points, and let $\mathfrak{L}$ be the set of all axis-parallel lines that intersect the grid. The number of lines in $\mathfrak{L}$ is $3S^2 = L$ lines. Each point in the grid is a joint for $\mathfrak{L}$, so there are S^3 joints. Therefore the number of joints is $\sim L^{3/2}$.

There is another example which has slightly more joints than this one. As a special case, consider the edges of a tetrahedron. A tetrahedron has six edges and four vertices. Each vertex lies in three non-coplanar edges, and so the tetrahedron gives a set of six lines with four joints. It is not hard to check that any six lines can form at most four joints. This tetrahedron example can be generalized to large numbers of lines in the following way.

Let $S \geq 3$ be a parameter. Consider S planes in $\mathbb{R}^3$ in general position. Any two of the planes intersect in a line. Let $\mathfrak{L}$ be this set of lines. The number of lines in $\mathfrak{L}$ is $L = \binom{S}{2}$ lines. Any three of the planes intersect in a point, and each such point is a joint of $\mathfrak{L}$. Therefore, $\mathfrak{L}$ has $\binom{S}{3}$ joints. If we take $S = 4$, we recover the tetrahedron example. In this example, the number of joints is still $\sim L^{3/2}$, but the constant is better than in the grid example.

This example is the best known example in the joints problem. As far as I know, it is possible that every set of $\binom{S}{2}$ lines in $\mathbb{R}^3$ determines at most $\binom{S}{3}$ joints. It is not known how to prove such a sharp estimate, but we will prove that these examples are sharp up to a constant factor.

Theorem 2.12. ([**GK1**], simplified by [**KSS**] and [**Q**]) Any L lines in space determine $\leq 10L^{3/2}$ joints.

Main Lemma. If $\mathfrak{L}$ is a set of lines in $\mathbb{R}^3$ that determines J joints, then one of the lines contains at most $3J^{1/3}$ joints.

Before proving the main lemma, let us see how it implies Theorem 2.12. Let $J(L)$ be the maximum number of joints that can be formed by L lines. If $\mathfrak{L}$ is a set of L lines, then the main lemma tells us that one of the lines contains at most $3J(L)^{1/3}$ of the joints. The number of joints not on this line is at most $J(L-1)$. This gives the following recursive bound for $J(L)$:

$$J(L) \leq J(L-1) + 3J(L)^{1/3}.$$

Iterating this argument, we see that

$$J(L) \leq J(L-1) + 3J(L)^{1/3} \leq J(L-2) + 2 \cdot 3J(L)^{1/3} \leq ... \leq L \cdot 3J(L)^{1/3}.$$

Now rearranging gives $J(L)^{2/3} \leq 3L$, which implies the theorem.

Now we turn to the proof of the main lemma.

Proof. Let P be a lowest degree non-zero polynomial that vanishes at every joint of $\mathfrak{L}$. By the parameter counting argument, Lemma 2.4, the degree of P is $\leq 3J^{1/3}$. We will prove the main lemma by contradiction, so we suppose that every line of $\mathfrak{L}$ contains $> 3J^{1/3}$ joints. By the vanishing lemma, Lemma 2.8, P must vanish on every line of $\mathfrak{L}$.

Next we study the gradient of P at each joint of $\mathfrak{L}$. Recall that for a smooth function $F : \mathbb{R}^n \to \mathbb{R}$, we write ∇F for the gradient $(\frac{\partial F}{\partial x_1}, ..., \frac{\partial F}{\partial x_n})$.

LEMMA 2.13. If x is a joint of $\mathfrak{L}$, and if a smooth function $F : \mathbb{R}^3 \to \mathbb{R}$ vanishes on the lines of $\mathfrak{L}$, then ∇F vanishes at x.

PROOF. By hypothesis, x lies in three non-coplanar lines of $\mathfrak{L}$. Let v_1, v_2, v_3 be tangent vectors for these three lines. The directional derivative of F in the direction v_i must vanish at x. So we have $\nabla F(x) \cdot v_i = 0$ for each i. Since the v_i are a basis of $\mathbb{R}^3$, we have $\nabla F(x) = 0$. □

So we see that the derivatives of P vanish at each joint. The derivatives have smaller degree than P. Since P was a minimal degree non-zero polynomial that vanishes at each joint, each derivative of P is identically zero! Then P must be constant. Since P is non-zero, it follows that there are no joints at all, and this gives a contradiction. □

Here is an example to illustrate the proof. Suppose that we start with an $A \times B \times C$ grid of points with $A < B < C$, and we let $\mathfrak{L}$ be the set of axis-parallel lines that intersect the grid. The number of lines in $\mathfrak{L}$ is $AB + AC + BC$, and the number of joints is ABC. All of the joints are contained in a union of A parallel planes. Therefore, there is a polynomial P of degree A which vanishes on all the joints (the polynomial is a product of linear factors, one for each plane). It is an exercise to check that P has minimal degree among all polynomials that vanish on the grid above. Moreover, up to scaling, P is the unique polynomial of degree A that vanishes on the grid. Each line of $\mathfrak{L}$ contains either A, B, or C joints, depending on which direction it is pointing. The polynomial P vanishes on all the lines of $\mathfrak{L}$ containing B joints and all the lines containing C joints, but on none of the lines with A joints. We see in this example that the minimal degree polynomial identifies the more important and less important lines. It locates at least one unimportant line with not too many joints on it.

2.6. Comments on the method

In each of the proofs in this chapter, there is a polynomial P that plays a crucial role in understanding the combinatorics. I think it's interesting to note that we prove that this crucial polynomial exists in a somewhat indirect way. One could imagine writing a formula for P in terms of the positions of the points in the set that we are studying. We will see an argument of this type later in Chapter 5. But in these arguments, we don't give a formula for P, and instead we prove that P must exist by a dimension-counting argument.

This aspect of the proofs reminds me of the probabilistic method. (The book *The Probabilistic Method*, [**AlSp**], by Alon and Spencer gives a clear, engaging exposition of this topic.) In the probabilistic method, one proves that an object with certain properties exists by considering a random object in some class, and showing that the random object has the given property with positive probability. Many of these arguments can be thought of as counting arguments. Suppose that we are interested in graphs with property X. There are $2^{\binom{n}{2}}$ different graphs on a set of n vertices. We try to count the number of graphs on n vertices that fail to have property X. If this number is smaller than $2^{\binom{n}{2}}$, then we know that there is a graph on n vertices with property X. A key insight of the probabilistic method

is that this approach may be much easier than constructing a particular graph and proving that it has property X.

For comparison, consider the following problem about polynomials. Suppose that we want to find a (non-zero) polynomial on $\mathbb{R}^2$ that vanishes at the points $(j, 2^j)$ for $j = 1, ..., 10^6$, and with the smallest possible degree. If we try to write down an explicit polynomial, we might come up with

$$P_1(x_1, x_2) = \prod_{j=1}^{10^6} (x_1 - j),$$

or

$$P_2(x_1, x_2) = \prod_{j=1}^{10^6} (x_2 - 2^j).$$

The polynomials P_1 and P_2 vanish on our set, and they each have degree 10^6. I think it is hard to find an explicit polynomial with degree much below 10^6. If one hasn't been shown the right way to think about the problem, then it may well seem that the minimal degree is on the order of 10^6.

But the parameter counting lemma, Lemma 2.4, tells us that there is a polynomial that vanishes on our set with degree at most 2000. This polynomial is probably very complicated to right down in any way. A polynomial in $\mathrm{Poly}_{2000}(\mathbb{R}^2)$ has over a million coefficients, and the parameter counting argument exploits all of them. Most of the polynomials in $\mathrm{Poly}_{2000}(\mathbb{R}^2)$ are very complicated to write down, but using linear algebra we can prove that there exists a polynomial with some desired properties. In the proofs in this chapter, we use this type of indirect method to prove that there is a polynomial with some desired properties, and then we exploit that polynomial to study combinatorial problems.

In the last few paragraphs, we discussed the parameter counting portion of the argument, but there is another basic issue about these proofs: why do we use polynomials at all? The statements of the problems don't involve polynomials, and the idea of using polynomials is a key insight in these proofs. I don't have a short clear answer for why polynomials play a special role in these problems. We will give a long discussion of this question in Chapter 3, and we will come back to it from time to time later in the book. Polynomials are a special class of functions. In these proofs, the main special property of polynomials that we use is the vanishing lemma. We do use some other properties in some of the arguments – for example, in the proof of the joints theorem, we use the fact that the derivative of a polynomial is a polynomial of lower degree. But the vanishing lemma is the most essential ingredient.

There are a lot of special properties of polynomials. Algebraic geometry could be described as the study of polynomials, or maybe as the study of the special properties of polynomials. There is no other class of functions which is at the center of such a large field of mathematics. Later in the book, we will use some other special properties of polynomials and more algebraic geometry.

The proofs in this chapter are models for all of the polynomial arguments in the book. In each proof, we use an indirect method, based on parameter counting, to find a polynomial with some desired properties. Then we exploit that polynomial to study our problem. In this second step, we bring into play the special properties of polynomials, especially the vanishing lemma or something related to it.

I think that it's surprising and interesting that this method has so many applications, both in combinatorics and in other fields. This method is sometimes called the polynomial method. Of course, there are a huge number of methods in mathematics involving polynomials. Maybe this type of argument should be called something like the polynomial/parameter counting/vanishing lemma method, but this name is too long. In any case, the aim of the book is to explore this type of argument.

2.7. Exercises

EXERCISE 2.1. Given a set of N points in $\mathbb{R}^3$, we proved that there is a non-zero polynomial of degree $\lesssim N^{1/3}$ that vanishes at all the points. Given any N lines in $\mathbb{R}^3$ prove that there is a non-zero polynomial of degree $\lesssim N^{1/2}$ that vanishes on all the lines.

State and prove a similar result for k-planes in $\mathbb{R}^n$ for any dimensions k, n.

EXERCISE 2.2. In this section, we discuss some alternate proofs of Lemma 2.5. We prove a stronger result called polynomial interpolation.

LEMMA 2.14. Suppose that $S = \{x_0, ..., x_D\}$ is a set of $D+1$ distinct points in the field $\mathbb{F}$. Define the evaluation map $E_S : \mathrm{Poly}_D(\mathbb{F}) \to \mathbb{F}^{D+1}$ by

$$E_S(Q) = (Q(x_0), ..., Q(x_D)).$$

The map E_S is an isomorphism.

Show that Lemma 2.14 implies Lemma 2.5.

The monomials $x^0, ..., x^D$ are a natural basis for $\mathrm{Poly}_D(\mathbb{F})$. Using this basis, we can write E_S as a matrix. This matrix is called the Vandermonde matrix. The Vandermonde matrix is the $N \times N$ matrix with (i, j) entry equal to x_j^i, where $0 \leq i, j \leq N$. (Here $N = D + 1$.)

One approach to prove Lemma 2.14 is to show that the determinant of the Vandermonde matrix is non-zero. The determinant is given by the following formula:

LEMMA 2.15. The determinant of the Vandermonde matrix is $\prod_{j_1 > j_2} (x_{j_1} - x_{j_2})$. In particular, if $x_0, ..., x_{N-1}$ are distinct, then the Vandermonde matrix is invertible.

There are several proofs in the literature of the Vandermonde determinant lemma. The most common proof uses Lemma 2.7, but there is also a proof using row reduction – cf. [**Va**].

Another approach to prove Lemma 2.14 is to check that E_S is surjective. To prove this, find a polynomial $f_j \in \mathrm{Poly}_D(\mathbb{F})$ so that $f_j(x_i) = 0$ if $i \neq j$ and $f_j(x_j) \neq 0$. This approach connects with the argument in Chapter 5.

EXERCISE 2.3. Prove the following result:

LEMMA 2.16. (Schwarz-Zippel lemma) Suppose that $A_i \subset \mathbb{F}$ are finite subsets, defined for $i = 1, ..., n$, with $|A_i| = N$ for all i. Suppose that $P \in \mathrm{Poly}_D(\mathbb{F}^n)$ is a non-zero polynomial. Prove that the number of zeroes of P in $A_1 \times ... \times A_n$ is at most DN^{n-1}.

An interesting special case is that $\mathbb{F} = \mathbb{F}_q$ is a finite field and $A_i = \mathbb{F}_q$ for all i.

EXERCISE 2.4. Suppose that $P \in \mathrm{Poly}(\mathbb{R}^n)$ is a non-zero polynomial. Prove that $Z(P)$, the zero set of P, has Lebesgue measure zero.

EXERCISE 2.5. We consider a collection of curves $\Gamma_a \subset \mathbb{F}_q^n$ parametrized by $a \in \mathbb{F}_q^{n-1}$. For each $a \in \mathbb{F}_q^{n-1}$, $1 \leq j \leq n-1$, suppose that $Q_{a,j} \in \mathrm{Poly}_d(\mathbb{F}_q)$. Let Γ_a be defined as the graph:

$$\Gamma_a := \{(Q_{a,1}(t), Q_{a,2}(t), ..., Q_{a,n-1}(t), t) \in \mathbb{F}_q^n | t \in \mathbb{F}_q\}.$$

Suppose also that $(Q_{a,1}(0), ..., Q_{a,n-1}(0)) = a$, so that $(a, 0) \in \Gamma_a$.

Prove that there is a constant $c(d, n) > 0$ so that

$$\left| \cup_{a \in \mathbb{F}_q^{n-1}} \Gamma_a \right| \geq c(d, n) q^n.$$

EXERCISE 2.6. The joints problem also makes sense in higher dimensions, and [**KSS**] and [**Q**] proved a generalization of Theorem 2.12 to all dimensions.

If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^n$, a joint of $\mathfrak{L}$ is defined to be a point that lies in n lines of $\mathfrak{L}$ pointing in linearly-independent directions.

THEOREM 2.17. A set of L lines in $\mathbb{R}^n$ determines at most $C_n L^{\frac{n}{n-1}}$ joints.

Prove this theorem.

Remark. The proof of the joints theorem in this chapter closely follows [**KSS**] and [**Q**]. It generalizes in a direct way to higher dimensions. The original proof in [**GK1**] was more complicated, and it did not generalize easily to higher dimensions.

EXERCISE 2.7. The axis-parallel case of the joints problem is already quite interesting. Suppose that $\mathfrak{L}_i$ is a set of lines in $\mathbb{R}^n$ parallel to the x_i axis. Let $\mathfrak{L} = \cup_i \mathfrak{L}_i$. A joint of $\mathfrak{L}$ is a point that lies in one line from each family $\mathfrak{L}_i$.

Loomis and Whitney proved [**LW**] that the number of joints of $\mathfrak{L}$ is at most $\prod_{i=1}^n |\mathfrak{L}_i|^{\frac{1}{n-1}}$. This implies that the number of joints is at most $|\mathfrak{L}|^{\frac{n}{n-1}}$.

Prove the axis-parallel case of the joints theorem. It is a good idea to start with the case $n = 3$.

The Loomis-Whitney theorem has interesting implications in analysis and geometry. We discuss some of these in Section 15.1.

EXERCISE 2.8. Suppose that $A, B, C \subset \mathbb{R}$ with $|A| < |B| < |C|$. Consider the grid $A \times B \times C \subset \mathbb{R}^3$. Let $P(x_1, x_2, x_3) = \prod_{a \in A}(x_1 - a)$. Prove that P is a minimal degree polynomial that vanishes on the grid. Moreover, prove that every minimal degree polynomial vanishing on the grid is a multiple of P.

What happens in $|A| = |B| = |C|$?

EXERCISE 2.9. This is a much harder exercise, or a project, based on similar ideas to the chapter.

Suppose that l_i is a line in $\mathbb{F}_q^n$ and suppose that $X_i \subset l_i$ is a subset with $|X_i| \geq q/2 = |l_i|/2$. Prove that

$$|\cup X_i| \geq c(d, n) \, |\cup l_i| \, .$$

(See [**NW**] for a proof and some generalizations.)

https://doi.org/10.1090//ulect/064/03

CHAPTER 3

Why polynomials?

The finite field Kakeya theorem and the joints theorem have short proofs using polynomials. At the current time, no one knows how to prove them without mentioning polynomials. Before people found the proofs from Chapter 2, they tried hard to attack these problems in other ways. It seems to be very difficult to prove these results without using the polynomial trick from Chapter 2. It would be interesting to understand why this is happening. The goal of the chapter is to start to explore this question.

I think it can be hard to appreciate the polynomial proofs of the finite field Kakeya and joints theorems without trying to prove them in other ways. So to start the chapter, we discuss finite field Kakeya and joints without polynomials. We explain some of the methods people have used to work on these problems and see the difficulties that they encounter.

After that, we discuss what properties of polynomials we used in the proofs from Chapter 2. What is special about polynomials that make them work well in these arguments? Could there be other spaces of functions that work equally well or better?

The real stars of this chapter are some interesting examples of sets of lines. The finite field Kakeya conjecture and the joints conjecture are true, but these examples show that some naive conjectures in a similar spirit are not true. Knowing about these examples shows that certain approaches cannot prove finite field Kakeya or joints.

These examples are all constructed using polynomials. Having interesting examples based on polynomials is another indication that polynomials play an important role in this circle of questions.

3.1. Finite field Kakeya without polynomials

In this section, we prove some estimates for the finite field Kakeya problem without the polynomial method. These estimates are all much weaker than Theorem 2.11.

Recall that a set $K \subset \mathbb{F}_q^n$ is called a Kakeya set if it contains a line in every direction. We want to study the minimal possible size of a Kakeya set $K \subset \mathbb{F}_q^n$. Theorem 2.11 gives the bound $|K| \gtrsim q^n$. Our first estimate gives a good answer for $n = 2$.

PROPOSITION 3.1. Suppose $s \leq q$. If $l_1, ..., l_s$ are lines in $\mathbb{F}_q^n$, then their union has cardinality at least $(1/2)qs$.

In particular, if $K \subset \mathbb{F}_q^n$ is a Kakeya set, then $|K| \geq (1/2)q^2$.

PROOF. We imagine adding the lines one at a time and keep track of the size of their union. The first line contains q points. The second must contain at least

$q-1$ points not in the first line. The third line must contain at least $q-2$ points not in the first two lines, etc. Therefore, the number of distinct points in the union of all s lines is at least $q+(q-1)+...+(q-s+1) > (1/2)qs$.

A Kakeya set always contains at least q lines, and so $|K| \geq (1/2)q^2$.

□

This estimate is very good when $n = 2$. For larger n, it is not such a good estimate. Examining the proof, we see that it only uses the fact that K contains at least q distinct lines. In dimension n, a Kakeya set contains at least q^{n-1} lines, so for $n \geq 3$ we have only used a small piece of the hypothesis.

Here is a second approach, called the bush method, which does better when n is large.

PROPOSITION 3.2. (Bush method) If $l_1, ..., l_M$ are lines in $\mathbb{F}_q^n$, then the number of points in their union is at least

$$(1/2)qM^{1/2}.$$

If $K \subset \mathbb{F}_q^n$ is a Kakeya set, then

$$|K| \geq (1/2)q^{\frac{n+1}{2}}.$$

PROOF. Let X be the union of the lines $l_1, ..., l_M$. Each of these lines contains q points of X. By the pigeonhole principle, there is a point $x \in X$ which lies in at least $qM|X|^{-1}$ of the lines l_i.

The set of lines l_i through the point x is called the bush of x. The lines in the bush of x are disjoint except at x, and their union lies in X. Therefore

$$(q-1)qM|X|^{-1} \leq |X|.$$

Rearranging this inequality, we get

$$|X| \geq (1/2)qM^{1/2}.$$

Now a Kakeya set $K \subset \mathbb{F}_q^n$ contains at least q^{n-1} lines. Plugging in $M = q^{n-1}$, we get $|K| \geq (1/2)q^{\frac{n+1}{2}}$. □

If $n \geq 4$, then the bush method gives a better estimate than our first method. If $n = 3$, both these methods give the estimate $|K| \gtrsim q^2$. Both methods actually apply to the union of any q^{n-1} distinct lines in $\mathbb{F}_q^n$. A plane contains more than q^2 distinct lines, and so we can find q^2 distinct lines in $\mathbb{F}_q^3$ whose union contains only q^2 points. But a Kakeya set $K \subset \mathbb{F}_q^3$ contains q^2 lines pointing in different directions. Only a small fraction of these lines can lie in a plane. To improve our estimate for the size of a Kakeya set in $\mathbb{F}_q^3$, we need to exploit this fact.

Here is a third approach to the finite field Kakeya problem called the hairbrush method. It combines ideas from the first two methods, and it takes advantage of the fact that not too many lines of a Kakeya set can lie in a plane.

PROPOSITION 3.3. (Hairbrush method, [**Wo1**]) Suppose that $l_1, ..., l_M$ are lines in $\mathbb{F}_q^n$, and suppose that at most $q+1$ of the lines lie in any plane. Then their union has cardinality at least

$$(1/3)q^{3/2}M^{1/2}.$$

If $K \subset \mathbb{F}_q^n$ is a Kakeya set, then $|K| \geq (1/2)|q|^{\frac{n+2}{2}}$.

PROOF. Let $X = \cup_i l_i$. If l_i is a line in K, then the hairbrush with stem l_i is defined to be the set of lines l_j which intersect l_i. (We don't include l_i itself in the hairbrush.) An average point of X lies in $qM|X|^{-1}$ lines l_i. If each point of X was about average, then each hairbrush would contain $\gtrsim q^2 M|X|^{-1}$ lines. We claim that there is always at least one hairbrush with $\geq (1/2)q^2 M|X|^{-1}$ lines. This claim is a simple counting argument, and we will come back to it below.

Fix a hairbrush containing $\geq (1/2)q^2 M|X|^{-1}$ lines. Let l_i be the stem of the hairbrush. Consider all of the 2-planes that contain the stem l_i. Each line in the hairbrush lies in exactly one of these 2-planes. If π is a 2-plane, then we let $H(\pi)$ denote the lines of the hairbrush that lie in π. Since every plane contains at most $q+1$ of our lines, we know that there are at most q lines in $H(\pi)$. Since each line in the hairbrush lies in exactly one plane π, we have

$$\sum_{\pi} |H(\pi)| \geq (1/2)q^2 M|X|^{-1}.$$

By our first method, Proposition 3.1, we know that the union of the lines in $H(\pi)$ contains at least $(1/3)|H(\pi)|q$ points of $\pi \setminus l_i$. As π varies among the 2-planes containing l_i, the sets $\pi \setminus l_i$ are disjoint. Therefore, the union of the lines in the hairbrush has cardinality at least

$$\sum_{\pi} (1/3)|H(\pi)|q \geq (1/6)q^3 M|X|^{-1}.$$

The lines in the hairbrush all lie in X, and so we get

$$(1/6)q^3 M|X|^{-1} \leq |X|.$$

Rearranging this gives the bound

$$|X| \geq (1/3)q^{3/2} M^{1/2}.$$

To finish this argument, we return to the simple claim about finding a hairbrush with many lines. Let $\mathfrak{L}$ be our set of lines $l_1, ..., l_M$. Let $\mu(p)$ be the number of lines of $\mathfrak{L}$ containing the point p. We know that $\sum_p \mu(p) = qM$. Consider triples of the form $(l, l', p) \in \mathfrak{L} \times \mathfrak{L} \times \mathbb{F}_q^n$, with $p \in l$ and $p \in l'$. The number of triples is $\sum_{p \in X} \mu(p)^2$. Since the function s^2 is convex in s, the number of triples is at least $|X| \left(qM|X|^{-1}\right)^2 = q^2 M^2 |X|^{-1}$. Since there are M lines in $\mathfrak{L}$, one of them must participate in at least $q^2 M|X|^{-1}$ triples. Call this line l_i. There are q "degenerate" triples of the form (l_i, l_i, p). But except for these, every triple (l_i, l_j, p) corresponds to a line l_j in the hairbrush of l_i, and every such line contributes exactly one triple. Therefore, the number of lines in the hairbrush of l_i is at least $q^2 M|X|^{-1} - q$. If $|X| \geq (1/2)qM$, then we are done, and otherwise, the number of lines in the hairbrush of l_i is at least $(1/2)q^2 M|X|^{-1}$ as desired.

This finishes the proof of the first part of the proposition. Now we apply this estimate to Kakeya sets. Suppose that $K = \cup l_i$ is a Kakeya set in $\mathbb{F}_q^n$, where the lines l_i point in different directions. We have $M \geq q^{n-1}$ lines l_i. Since they point in different directions, there are at most $q+1$ of them in any plane. Plugging this into our first bound, we see that $|K| \geq (1/3)q^{\frac{n+2}{2}}$.

□

In $\mathbb{F}_q^3$, the hairbrush method shows that a Kakeya set has size $\gtrsim q^{5/2}$. More generally, it shows that if $\{l_i\}$ is a set of q^2 lines in $\mathbb{F}_q^3$ with at most q lines in any

2-plane, then $|\cup_i l_i| \gtrsim q^{5/2}$. If q is a square, then there is a remarkable example which shows that this estimate is tight. This example is called a Hermitian variety $H \subset \mathbb{F}_q^3$. It is an algebraic surface in $\mathbb{F}_q^3$, it has $|H| \sim q^{5/2}$, and it contains a set of $\sim q^2$ lines $\mathfrak{L}_H$ with far less than q lines of $\mathfrak{L}_H$ in any 2-plane. In the next section, we describe this example.

(If q is prime, then Ellenberg and Hablicsek [**ElHa**] recently proved that this type of example cannot occur. If $\{l_i\}$ is a set of q^2 lines in $\mathbb{F}_q^3$ with at most q lines in any 2-plane, and if q is prime, then they proved that $|\cup_i l_i| \gtrsim q^3$. Their proof uses the polynomial method. In particular, it builds on the ideas from Chapter 11.)

Hermitian varieties were described by Bose in [**BC**] in the 1960's as an interesting example in algebraic geometry. The connection between this example and the Kakeya problem was explained by Katz, Laba, and Tao in [**KLT**]. They used a closely related example called the Heisenberg group. The paper [**KLT**] describes the Heisenberg group in $\mathbb{C}^3$ and the finite field version appears in [**MT**].

The Hermitian variety example shows that in order to improve the bound $q^{5/2}$ for a Kakeya set in $\mathbb{F}_q^3$, we cannot just use the fact that there are not too many lines of a Kakeya set in any 2-plane. We need to get some additional mileage out of the fact that the lines point in different directions, and it is quite challenging to do this. Before the polynomial method, the best proven lower bound for the size of a Kakeya set in $\mathbb{F}_q^3$ was $\sim q^{5/2}$.

There was some further interesting progress on the Kakeya problem using combinatorial number theory. If q is prime, a slightly stronger estimate for the size of Kakeya sets in $\mathbb{F}_q^3$ was proven in [**BKT**]. For large dimensions n, combinatorial number theory ideas led to stronger estimates than the hairbrush argument, but they were still much weaker than Theorem 2.11. They had the form $|K| \gtrsim q^{\alpha n}$ for some $\alpha > 1/2$. The best known value of α was a little less than .6. For an introduction to this interesting line of attack, see [**Lab**] and [**Ta1**].

We included this discussion to help put the finite field Kakeya problem in perspective. The first couple methods show how a reasonable person might start to think about the problem and the hairbrush argument is a development of those methods. It is remarkable that one can prove much better estimates using polynomials.

3.2. The Hermitian variety

The Hermitian variety is an interesting variety $H \subset \mathbb{F}_q^3$ which we can define whenever q is a square. In this section, we discuss the case that $q = p^2$ where p is a prime. Our main interest is in large primes p, and in particular we will assume that p is odd. The Hermitian variety $H \subset \mathbb{F}_q^3$ is defined by the equation

$$x_1^{p+1} + x_2^{p+1} + x_3^{p+1} = 1.$$

The Hermitian variety contains a lot of lines that intersect each other in a complicated way. From the point of view of combinatorics, it is probably the most interesting configuration of lines that we know about.

Proposition 3.4. The variety H contains $\sim q^{5/2}$ points of $\mathbb{F}_q^3$. The variety H contains $\sim q^2$ lines. We denote the set of lines in H by $\mathfrak{L}_H$. Any plane contains $\lesssim q^{1/2}$ lines of $\mathfrak{L}_H$.

This example was discovered by Bose and Chakravarti [**BC**], and independently by Segre [**Seg**] in the 1960s. Mockenhaupt and Tao rediscovered a small variation of this example, [**MT**]. (Steve Kleiman recently pointed out the work of [**BC**] and [**Seg**] to the combinatorics community.)

We begin with some basic facts about the multiplicative group $\mathbb{F}_q^*$. The group $\mathbb{F}_q^*$ is a cyclic group of order $q-1 = (p+1)(p-1)$ (cf. Theorem 1.9 in Chapter 4 of [**Lan**]). Now $\mathbb{F}_p^* \subset \mathbb{F}_q^*$ is a cyclic subgroup of order $p-1$. Therefore, we can characterize $\mathbb{F}_p^* \subset \mathbb{F}_q^*$ in the following way:

$$\mathbb{F}_p^* = \{y \in \mathbb{F}_q^* | y^{p-1} = 1\}.$$

This observation leads to some special properties of the operation $x \mapsto x^{p+1}$:

LEMMA 3.5. For every $x \in \mathbb{F}_q^*$, $x^{p+1} \in \mathbb{F}_p^*$. On the other hand, for every $y \in \mathbb{F}_p^*$, the equation $x^{p+1} = y$ has exactly $p+1$ solutions $x \in \mathbb{F}_q^*$.

PROOF. Since $\mathbb{F}_q^*$ is cycle of order $q-1 = (p-1)(p+1)$, we see that $(x^{p+1})^{p-1} = x^{q-1} = 1$, and so $x^{p+1} \in \mathbb{F}_p^*$. On the other hand, since $\mathbb{F}_q^*$ is cyclic of order $(p-1)(p+1)$, the second claim follows. □

As a corollary, we can estimate the size of H.

COROLLARY 3.6. $|H| \sim q^{5/2}$.

PROOF SKETCH. For any $x_1, x_2 \in \mathbb{F}_q$, we study the set of x_3 so that $(x_1, x_2, x_3) \in H$. We let $y = 1 - x_1^{p+1} - x_2^{p+1} \in \mathbb{F}_p$. If $y = 0$, then $(x_1, x_2, x_3) \in H$ if and only if $x_3 = 0$. But if $y \neq 0$, then Lemma 3.5 says that there are $p+1 \sim q^{1/2}$ values of x_3 so that $(x_1, x_2, x_3) \in H$. It's straightforward to check that for the majority of $(x_1, x_2) \in \mathbb{F}_q^2$, $y \neq 0$, and so $|H| \sim q^{1/2} q^2 = q^{5/2}$. □

Before we describe the lines in H, we introduce a piece of algebraic structure that helps to understand H better. We define a conjugation operation on $\mathbb{F}_q$: for any $x \in \mathbb{F}_q$, $\bar{x} := x^p$. This conjugation is an involution because $(x^p)^p = x^q = x$. In particular $\bar{x} = 0$ if and only if $x = 0$.

This conjugation is analogous to complex conjugation. In fact, if p is congruent to 3 mod 4, then there is no solution to the equation $y^2 = -1$ with $y \in \mathbb{F}_p$, and we can define $\mathbb{F}_q = \mathbb{F}_p[i]$. In this case, if $x = y_1 + iy_2$, then the reader can check that $\bar{x} := x^q = y_1 - iy_2$. We won't need this in the sequel, but it might be helpful to keep it in mind.

Using our definition of conjugation, $\bar{x} = x^p$, the equation defining H can be rewritten in the form

$$x_1\bar{x}_1 + x_2\bar{x}_2 + x_3\bar{x}_3 = 1.$$

The left-hand side is reminiscent of a Hermitian inner product. In particular, the Hermitian variety is a finite-field analogue of the unit sphere in $\mathbb{C}^3$, defined by the equations $z_1\bar{z}_1 + z_2\bar{z}_2 + z_3\bar{z}_3 = 1$. If v and w are vectors in $\mathbb{F}_q^n$, we will write

$$v \cdot \bar{w} = \sum_{i=1}^{n} v_i \bar{w}_i.$$

We observe a couple simple algebraic facts about this expression, analogous to standard facts about a Hermitian inner product in $\mathbb{C}^n$.

For any $v \in \mathbb{F}_q^n$, $v \cdot \bar{v} \in \mathbb{F}_p$. Indeed, $v \cdot \bar{v} = \sum_i v_i \bar{v}_i = \sum_i v_i^{p+1}$. And for each i, $v_i^{p+1} \in \mathbb{F}_p$ by Lemma 3.5.

Also,

$$\overline{(v \cdot \bar{w})} = \bar{v} \cdot w.$$

Because we are working in characteristic p, $(\sum_i a_i)^p = \sum_i a_i^p$. Therefore, the left-hand side of the last equation is

$$\left(\sum_i v_i w_i^p\right)^p = \sum_i v_i^p w_i^q = \sum_i v_i^p w_i = \bar{v} \cdot w.$$

In particular, $v \cdot \bar{w} = 0$ if and only if $\bar{v} \cdot w = 0$.

With these tools we now study the lines in H. For any $a_1, b_1, a_2, b_2 \in \mathbb{F}_q$, we let $l(a, b)$ be the line given by the equations

$$x_1 = a_1 x_3 + b_1; x_2 = a_2 x_3 + b_2.$$

There are q^4 lines $l(a, b)$ and these are almost all of the lines in $\mathbb{F}_q^3$. We want to figure out which lines $l(a, b)$ are contained in H. To see if $l(a, b)$ is contained in H, we plug in the equations $x_i = a_i x_3 + b_i$ into the formula $x_1^{p+1} + x_2^{p+1} + x_3^{p+1} = 1$. Grouping the terms by powers of x_3, we see that $l(a, b)$ lies in H if and only if

$$(a_1^{p+1} + a_2^{p+1} + 1)x_3^{p+1} + (a_1^p b_1 + a_2^p b_2)x_3^p + (b_1^p a_1 + b_2^p a_2)x_3 + (b_1^{p+1} + b_2^{p+1} - 1) = 0$$
$$\text{for all } x_3 \in \mathbb{F}_q.$$

The left-hand side is a polynomial of degree $p + 1 < q$, and so by the vanishing lemma it can vanish at all q points $x_3 \in \mathbb{F}_q$ if and only if all the coefficients vanish. So we see that $l(a, b) \subset H$ if and only if the following four equations hold:

(1) $a_1^p b_1 + a_2^p b_2 = 0$.
(2) $b_1^p a_1 + b_2^p a_2 = 0$.
(3) $a_1^{p+1} + a_2^{p+1} = -1$.
(4) $b_1^{p+1} + b_2^{p+1} = 1$.

These equations become much simpler if we rewrite them using our conjugation notation and using the vectors $a = (a_1, a_2)$ and $b = (b_1, b_2)$. The first two equations become $\bar{a} \cdot b = 0$ and $a \cdot \bar{b} = 0$. As we discussed above, these two equations are equivalent! The list of equations becomes

(1) $a \cdot \bar{b} = 0$
(2) $a \cdot \bar{a} = -1$.
(3) $b \cdot \bar{b} = 1$.

Writing the equations in this way they become much more approachable, and we can get a good estimate for $|\mathfrak{L}_H|$. The first equation has $\sim q^3$ solutions, because for any non-zero vector $b \in \mathbb{F}_q^2$, there are exactly q values of a that solve the equation. For most of these solutions $a \cdot \bar{a}$ and $b \cdot \bar{b}$ are non-zero. In other words, if we define

$$S := \{a, b \text{ so that } a \cdot \bar{b} = 0, a \cdot \bar{a} \neq 0, b \cdot \bar{b} \neq 0\},$$

then $|S| \sim q^3$.

Now on S, the values of $(a \cdot \bar{a}, b \cdot \bar{b})$ are evenly distributed within $(\mathbb{F}_p^*)^2$. This follows by a symmetry argument. Lemma 3.5 implies that as λ varies in $\mathbb{F}_q^*$, $\lambda\bar{\lambda}$ is evenly distributed in $\mathbb{F}_p^*$. Now $(\mathbb{F}_q^*)^2$ acts on S: for $(\lambda, \mu) \in (\mathbb{F}_q^*)^2$, the action sends (a, b) to $(\lambda a, \mu b)$. This action transforms $(a \cdot \bar{a}, b \cdot \bar{b})$ to $(\lambda\bar{\lambda} a \cdot \bar{a}, \mu\bar{\mu} b \cdot \bar{b})$, and so $(a \cdot \bar{a}, b \cdot \bar{b})$ must be evenly distributed. Therefore,

$$|\mathfrak{L}_H| = (p - 1)^{-2}|S| \sim q^2.$$

Our description of $\mathfrak{L}_H$ also allows us to see that $\mathfrak{L}_H$ is not a Kakeya set of lines. The direction of the line $l(a,b)$ is determined by a, and the vector a has to satisfy the equation $a_1\bar{a}_1 + a_2\bar{a}_2 = -1$. This equation has $\sim pq \sim q^{3/2}$ solutions. So the lines of $\mathfrak{L}_H$ only point in $\sim q^{3/2}$ different directions, and for each direction there are $\sim q^{1/2}$ parallel lines in that direction.

We have now shown that $|\mathfrak{L}_H| \sim q^2$, the hardest and most interesting part of Proposition 3.4. We want to briefly sketch an alternative way to estimate $|\mathfrak{L}_H|$, based on the symmetries of H. The set H is very symmetrical: there is a transitive group of symmetries, and so every point of H is equivalent to any other point. This symmetry group is an analogue of the unitary group. We define

$$\mathrm{U}(\mathbb{F}_q^n) := \{g \in \mathrm{GL}(\mathbb{F}_q^n) \text{ so that } v \cdot \bar{w} = (gv) \cdot \overline{(gw)} \text{ for all } v, w \in \mathbb{F}_q^n\}.$$

Exercise 3.1. The group $\mathrm{U}(\mathbb{F}_q^3)$ acts transitively on the Hermitian variety H. This is the $\mathbb{F}_q$-analogue of the fact that the standard unitary group $U(3)$ acts transitively on the unit sphere in $\mathbb{C}^3$ defined by $z_1\bar{z}_1 + z_2\bar{z}_2 + z_3\bar{z}_3 = 1$.

Given this transitive symmetry, it is easier to estimate $|\mathfrak{L}_H|$, because we only have to estimate the number of lines through a single convenient point. We consider the point $(1,0,0) \in H$, and we will construct $p+1$ lines of $\mathfrak{L}_H$ through this point. The lines we construct all lie in the plane $x_1 = 1$. The intersection of H with this plane is given by the equation

$$x_2^{p+1} = -x_3^{p+1}.$$

This intersection contains the line of the form $x_1 = 1; x_2 = ax_3$ whenever $a^{p+1} = -1$. There are $p+1$ such values of a. All of these lines go through the point $(1,0,0)$. In summary, there are $p+1$ lines of $\mathfrak{L}_H$ through the point $(1,0,0)$.

Using the symmetry under the unitary group, we see that there are at least $p+1$ lines of $\mathfrak{L}_H$ through every point of H. Since each line of $\mathfrak{L}_H$ contains q points of H, and each point of H lies in at least $p+1$ lines of $\mathfrak{L}_H$, we get

$$q|\mathfrak{L}_H| \geq (p+1)|H| \sim q^3.$$

This argument gives an alternate derivation of the fact that $|\mathfrak{L}_H| \sim q^2$.

As an aside, it turns out that the $p+1$ lines that we constructed are all of the lines of $\mathfrak{L}_H$ through the point $(1,0,0)$. All these lines lay in the plane $x_1 = 1$. Using the symmetry of H, it follows that the lines of $\mathfrak{L}_H$ do not form any joints at all: at every point $x \in H$, all the lines of $\mathfrak{L}_H$ containing x lie in a plane.

Finally, we have to check that any plane contains $\lesssim q^{1/2}$ lines of $\mathfrak{L}$. To prove this, we will have to understand a little about the condition $v \cdot \bar{v} = 0$. Unlike in $\mathbb{C}^n$, there can be non-zero vectors $v \in \mathbb{F}_q^n$ with $v \cdot \bar{v} = 0$. To see this, recall that for any $y \in \mathbb{F}_p^*$, there exists some $x \in \mathbb{F}_q^*$ with $x\bar{x} = x^{p+1} = y$. So we can choose $v_1 \in \mathbb{F}_q^*$ with $v_1\bar{v}_1 = 1$ and $v_2 \in \mathbb{F}_q^*$ with $v_2\bar{v}_2 = -1$, and then $v \cdot \bar{v} = 0$. However, in $\mathbb{F}_q^3$, the set of vectors v with $v \cdot \bar{v} = 0$ does not contain a whole 2 plane. We state this as a lemma.

Lemma 3.7. Suppose that $V \subset \mathbb{F}_q^3$ is a 2-dimensional subspace. Then there exists a vector $v \in V$ with $v \cdot \bar{v} \neq 0$.

Proof. For any $w \in \mathbb{F}_q^3$, define the linear map $L_w : \mathbb{F}_q^3 \to \mathbb{F}_q$ by $L_w(v) = v \cdot \bar{w}$. In coordinates, we have

$$L_w(v) = \bar{w}_1 v_1 + \bar{w}_2 v_2 + \bar{w}_3 v_3.$$

If $w \neq 0$, then L_w is a 2-dimensional subspace of $\mathbb{F}_q^3$. If $\operatorname{Ker} L_w = \operatorname{Ker} L_{w'}$ then we must have $\bar{w} = \lambda \bar{w}'$ for some $\lambda \in \mathbb{F}_q^*$. Conjugating this equation, we see that $w = \bar{\lambda} w'$.

Since V is 2-dimensional, we can choose $w, w' \in V$ so that $\operatorname{Ker} L_w \neq \operatorname{Ker} L_{w'}$. At least one of these kernels must be different from V. After relabelling, we can assume that $\operatorname{Ker} L_w \neq V$, and so we can choose $u \in V$ with $u \cdot \bar{w} \neq 0$. After multiplying u by a scalar, we can assume that $u \cdot \bar{w} = 1$. Conjugating, we also see that $\bar{u} \cdot w = 1$.

We use this information to expand $(u+w) \cdot \overline{(u+w)}$:

$$(u+w)\overline{(u+w)} = u \cdot \bar{u} + w \cdot \bar{w} + 2.$$

Now, u, w, and $u+w$ all lie in V. We assume that p is odd and so 2 is non-zero, and so one of the expression $u \cdot \bar{u}$, $w \cdot \bar{w}$, or $(u+w)\overline{(u+w)}$ must be non-zero. □

Now we are ready to check that H does not contain a 2-plane.

Lemma 3.8. The set $H \subset \mathbb{F}_q^3$ does not contain a 2-plane.

Proof. We can write an arbitrary 2-plane π in the form $a + V$, where $a \in \mathbb{F}_q^3$ and V is a 2-dimensional subspace of $\mathbb{F}_q^3$. By the last lemma, there is a vector $v \in V$ with $v \cdot \bar{v} \neq 0$. Consider the line $l \subset \pi$ parametrized by $t \mapsto a + tv$. We will check that the line l is not contained in H, and so the plane π is not contained in H.

If l were contained in H, we would have

$$(a+tv)\overline{(a+tv)} - 1 = 0 \text{ for all } t \in \mathbb{F}_q.$$

Using that we are in characteristic p, the left-hand side is

$$(a+tv)\cdot(a+tv)^p - 1 = (a+tv)\cdot(\bar{a}+t^p\bar{v}) - 1 = (v\cdot\bar{v})t^{p+1} + (a\cdot\bar{v})t^p + (v\cdot\bar{a})t + (a\cdot\bar{a}-1).$$

This expression is a polynomial in t. Since $v \cdot \bar{v} \neq 0$, it is a non-zero polynomial of degree $p+1 < q$. By the vanishing lemma, Lemma 2.8, it cannot vanish at all q values of t. Therefore, the line l is not contained in H. □

Now we can estimate the number of lines in $H \cap \pi$ for a 2-plane π. Since H does not contain π, the restriction of the polynomial $x_1^{p+1} + x_2^{p+1} + x_3^{p+1} - 1$ to the plane π must be non-zero. It has degree at most $p+1$, and so by the Schwarz-Zippel lemma (see Exercise 2.3), $|H \cap \pi| \leq (p+1)q \lesssim q^{3/2}$.

We can now bound the number of lines in $H \cap \pi$ using Proposition 3.1. If $H \cap \pi$ contains $s \leq q/2$ lines, then Proposition 3.1 says that $|H \cap \pi| \geq (1/2)qs$. Since $|H \cap \pi| \lesssim q^{3/2}$, we conclude that $s \lesssim q^{1/2}$, and so there are $\lesssim q^{1/2}$ lines of $\mathfrak{L}_H$ in any plane. This finishes the proof of Proposition 3.4.

Bose and Chakravarti [**BC**] also studied higher-dimensional Hermitian varieties, $H_n \subset \mathbb{F}_q^n$ defined by the equation $\sum_{i=1}^n x_i \bar{x}_i = 1$. They showed that higher dimensional Hermitian varieties contain many k-planes for k on the order of $n/2$.

Exercise 3.2. Let $Q \subset \mathbb{F}_q^4$ be the degree 2 hypersurface defined by the equation $x_1^2 + x_2^2 - x_3^2 - x_4^2 = 1$. Prove that each point $x \in Q$ lies in $\sim q$ lines in Q. Also check that Q contains $\sim q^3$ points and $\sim q^3$ lines. Check that the lines through each point $x \in Q$ lie in a 3-plane.

3.3. Joints without polynomials

The joints problem has a short proof with high degree polynomials, but it seems hard to prove without polynomials. This is somewhat surprising because the statement of the problem involves only points and lines and planes. Why is it hard to prove the joints theorem just mentioning linear objects?

In this section, we consider an approach to the joints problem just mentioning lines and planes, and we see why it leads to very weak bounds.

A joint is a triple intersection point (a point where ≥ 3 lines meet). So we begin by asking how many triple intersection points can be formed by L lines. Since two lines intersect in at most one point, the number of triple intersection points is $\lesssim L^2$. There is an example where the number of triple intersection points is $\sim L^2$. The example is a grid of horizontal, vertical, and diagonal lines in $\mathbb{R}^2$. Let $\mathfrak{L}$ denote the following set of lines:

- Horizontal lines $y = b$ for each integer $b = 1, ..., N$.
- Vertical lines $x = a$ for each integer $a = 1, ..., N$.
- Diagonal lines $x - y = c$ for each integer $c = -N, ..., N$.

This set has $L = 4N + 1$ lines, and it has a triple intersection point at each integer point (a, b) with $1 \leq a, b \leq N$. We let E_0 denote this integer grid of $\sim L^2$ triple intersection points.

The triple intersection points of $\mathfrak{L}$ are not joints, because all the lines lie in a plane. We can think of $\mathbb{R}^2 \subset \mathbb{R}^3$, so that the lines lie in $\mathbb{R}^3$. Then we may ask if this configuration is flexible? Can we perturb the lines of $\mathfrak{L}$, preserving all the triple intersections, and making the lines not coplanar?

Let us make the question more precise. If $\mathfrak{L} = \{l_1, l_2, ...\}$ is a set of lines and $E = \{p_1, p_2, ...\}$ is a set of points, then the incidence matrix $I(\mathfrak{L}, E)$ is the matrix with (i, j) entry equal to 1 if l_i contains p_j and equal to zero if l_i doesn't contain p_j. We call $(\mathfrak{L}', E')$ a perturbation of $(\mathfrak{L}, E)$ if they have the same incidence matrix. In particular, if $(\mathfrak{L}', E_0')$ is a perturbation of our first example $(\mathfrak{L}, E_0)$, then each point of E_0' is a triple point of $\mathfrak{L}'$. Now we can ask a precise question about perturbing $(\mathfrak{L}, E_0)$. Is there a perturbation of $(\mathfrak{L}, E_0)$ so that most of the points of E_0' are joints of $\mathfrak{L}'$?

The answer to this questions is no. We sketch a proof, using only lines and planes. A triangle is defined to be a set of three lines and three points so that each line contains exactly two of the points. A triangle in $(\mathfrak{L}, E_0)$ is defined to be a triangle where the three lines are in $\mathfrak{L}$ and the three points are in E_0. Triangles are preserved by perturbation: each triangle of $(\mathfrak{L}, E_0)$ corresponds to a triangle of $(\mathfrak{L}', E_0')$. Any triangle lies in a unique plane. If (l_1, l_2, l_3) are three lines of a triangle, lying in a plane π, and if l is a fourth line that intersects two lines of the triangle at distinct points, then l must also lie in the plane π. Using this fact, we can force more and more lines to lie in a plane.

Let T be the triangle with edges $x = 1$, $y = N$, and $x - y = 0$ (in $\mathfrak{L}$), and vertices $(1, 1), (1, N), (N, N)$ in E_0. Many lines of $\mathfrak{L}$ intersect T at two distinct points of E_0. All of these lines lie in the same plane as T. In the perturbation $(\mathfrak{L}', E_0')$, there is a corresponding triangle T'. It lies in a plane π'. Many lines of $\mathfrak{L}'$ intersect T' at two distinct points of E_0'. All these lines must also lie in π'. With a little more work, we see that almost all the lines of $\mathfrak{L}'$ lie in the plane π'. (There are only two exceptions: two of the diagonal lines of $\mathfrak{L}$ contain exactly one point of

E_0. The corresponding lines of $\mathfrak{L}'$ don't have to lie in π'. So $\mathfrak{L}'$ can determine at most two joints.)

We call this argument the triangle method. The triangle method only mentions points, lines, and planes, and it seems like a reasonable approach to the joints problem. The paper [**GS**] uses the triangle method to prove a very weak estimate on the number of joints: L lines determine $o(L^2)$ joints. Let us describe why it is hard to prove a good estimate using the triangle method.

Let $\mathfrak{L}$ be the same grid of horizontal, vertical, and diagonal lines as above, and suppose that E is a subset of E_0. For any subset $E \subset E_0$, we can ask whether there is a perturbation $(\mathfrak{L}', E')$ so that the points of E' are all joints of $\mathfrak{L}'$. If the pair $(\mathfrak{L}, E)$ does not contain any triangles, then the triangle method does not give any information about $(\mathfrak{L}', E')$. For example, Figure 3.1 is a picture of a triangle-free set $E \subset E_0$.

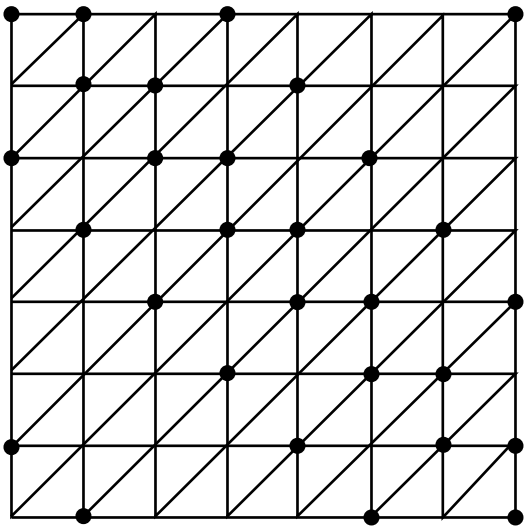

FIGURE 3.1. A triangle-free set.

The dots are the points of E. In Figure 3.1, is it possible to perturb the lines of $\mathfrak{L}$, preserving all the triple intersections in the set E, and converting the points of E into joints?

This type of example leads to the following question: how large can E be if $(\mathfrak{L}, E)$ does not contain any triangles? The paper [**GS**] shows that $|E| = o(L^2)$. But triangle-free subsets can be suprisingly large.

THEOREM 3.9. (Ajtai and Szemerédi, [**AjSz**]) For any $\varepsilon > 0$, for all L sufficiently large, there is a subset $E \subset E_0$ so that $(\mathfrak{L}, E)$ contains no triangles and yet $|E| \gtrsim L^{2-\varepsilon}$.

In the example of Theorem 3.9, is it possible to perturb the lines of $\mathfrak{L}$, preserving all the triple intersections in the set E, and converting the points of E into joints? Because of the joints theorem, we know that this is impossible. But without mentioning polynomials, it seems hard to rule out this possibility. See [**GS**] for further discussion.

To finish this section, we construct the set E. As in the example in Figure 3.1, the set E will be a union of diagonals of slope -1. In other words, for a well-chosen subset $B \subset [1, ..., 2N]$, the set E is given by

$$E := \{(x, y) \in [1, ..., N]^2 \text{ so that } x + y \in B\}.$$

Triangles in E correspond to 3-term arithmetic progressions in B. Recall that a 3-term arithmetic progression is a sequence of the form $a, a + d, a + 2d$.

LEMMA 3.10. If the set B contains no 3-term arithmetic progressions, then the set E contains no triangles of $(\mathfrak{L}, E)$.

This is easiest to see in a picture. See Figure 3.2.

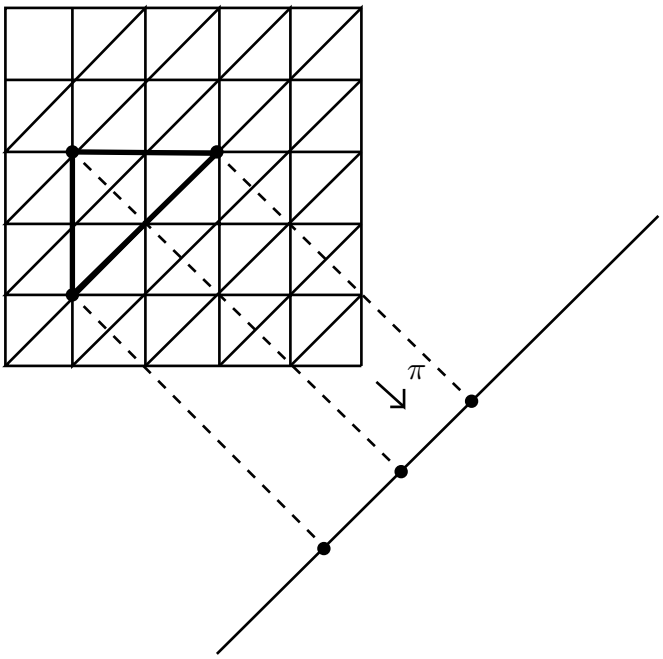

FIGURE 3.2. A triangle in E leads to an arithmetic progression B.

PROOF. Suppose that the lines l_1, l_2, l_3 form a triangle of $(\mathfrak{L}, E)$. Every pair of lines must intersect in a point, so no two of them are parellel, and so there must be one horizontal line, one vertical line, and one diagonal line. We label them so that l_1 is horizontal, l_2 is diagonal, and l_3 is vertical. Let $x_1 = (a_1, b_1)$ be the intersection of l_2 with l_3, and $x_2 = (a_2, b_2)$ be the intersection of l_1 and l_3 and $x_3 = (a_3, b_3)$ be the intersection of l_1 and l_2. We have $x_1, x_2, x_3 \in E$, and so $a_1 + b_1, a_2 + b_2, a_3 + b_3 \in B$. But we claim that the geometry of the situation forces $a_1 + b_1, a_2 + b_2, a_3 + b_3$ to be a 3-term arithmetic progression. This contradiction will prove the lemma.

We give an algebraic proof as follows. The points x_1, x_2 are on the same vertical line l_3 and so $a_1 = a_2$. Next the points x_2, x_3 are on the same horizontal line l_1, and so $b_2 = b_3$. Finally, the points x_1, x_3 are on the same diagonal line, and so

$a_1 - b_1 = a_3 - b_3$. Using these equations, we want to check that $a_1+b_1, a_2+b_2, a_3+b_3$ form a 3-term arithmetic progression. This boils down to checking

$$[a_3 + b_3] - [a_2 + b_2] = [a_2 + b_2] - [a_1 + b_1].$$

Using the equations:

$$[a_3 + b_3] - [a_2 + b_2] = a_3 - a_2 = a_3 - a_1 = b_3 - b_1 = b_2 - b_1 = [a_2 + b_2] - [a_1 + b_1].$$

□

This leads to the question, what is the largest possible size of a subset $B \subset [-N, ..., N]$ with no 3-term arithmetic progression? Behrend constructed a surprisingly large example of such a set.

THEOREM 3.11. (Behrend) For any $\varepsilon > 0$, for all N sufficiently large, there is a subset $B \subset [-N, ..., N]$ with no 3-term arithmetic progression and $|B| \gtrsim N^{1-\varepsilon}$.

Combining Theorem 3.11 and Lemma 3.10 gives Theorem 3.9.

The definition of an arithmetic progression makes sense in any abelian group. In particular, it makes sense in $\mathbb{Z}^n$ for any dimension n. Behrend's construction begins by finding a large subset A of a high-dimensional cube $[-S, S]^n$, and then "transferring" A to a subset $B \subset [-N, ..., N]$.

LEMMA 3.12. For any dimension n, for any $S \geq 1$, there is a subset $A \subset [-S, ..., S]^n$ with no 3-term arithmetic progression, and with

$$|A| \geq c(n)S^{n-2}.$$

PROOF. We use coordinates $x_0, ..., x_{n-1}$ on $\mathbb{R}^n$. The set A will be the set of lattice points on the sphere $\sum_{i=0}^{n-1} x_i^2 = M$ for a well-chosen M. There are more than S^n points in the cube $[-S, S]^n$. For each such point, $\sum_{i=0}^{n-1} x_i^2$ is an integer in the range $0, ..., nS^2$. By the pigeonhole principle, we can choose a value of M so that the sphere $\sum_{i=0}^{n-1} x_i^2 = M$ contains $\geq c(n)S^{n-2}$ points of the cube $[-S, ..., S]^n$. We let A be this set of points.

The points of a 3-term arithmetic progression $a, a + d, a + 2d \in \mathbb{R}^n$ all lie on a line. By convexity, a line intersects a sphere in at most two points, and so a sphere does not contain any 3-term arithmetic progression. □

Now we describe how to use A to construct our subset $B \subset [-N, ..., N]$. Suppose that $x = (x_0, ..., x_{n-1}) \in [-S, ..., S]^n$. Define

$$\phi(x) = \sum_{i=0}^{n-1} (10S)^i x_i.$$

We define $N = (10S)^n$ and we define B to be the image $\phi(A)$. We claim that ϕ is injective and that B contains no 3-term arithmetic progression. Since ϕ is injective,

$$|B| \geq c(n)S^{n-2} \geq c(n)N^{\frac{n-2}{n}}.$$

For any $\varepsilon > 0$, we choose n so that $1 - \varepsilon \leq \frac{n-2}{n}$, and we get Behrend sets with $|B| \geq c(\varepsilon)N^{1-\varepsilon}$ as desired. It remains to check that ϕ is injective and that B contains no 3-term arithmetic progression.

To check that ϕ is injective, suppose that $\phi(x) = \phi(y)$. In other words,

$$\sum_{i=0}^{n-1}(10S)^i x_i = \sum_{i=0}^{n-1}(10S)^i y_i.$$

We have to check that $x_i = y_i$ for all i. We first claim that $x_{n-1} = y_{n-1}$. Bringing the x_{n-1} and y_{n-1} to one side and all other terms to the other side, we get

$$(10S)^{n-1}(x_{n-1} - y_{n-1}) = \sum_{i=0}^{n-2}(10S)^i(y_i - x_i).$$

Since $x_i, y_i \in [-S, ..., S]$, the right hand side is at most $2 \cdot (10S)^{n-2}(2S) < (10S)^{n-1}$. Therefore, $|x_{n-1} - y_{n-1}| < 1$. Since $x_{n-1}, y_{n-1} \in \mathbb{Z}$, it follows that $x_{n-1} = y_{n-1}$. By the same argument, we can show that $x_j = y_j$ for all $j \leq n-1$ by backwards induction on j. Suppose that $x_i = y_i$ for all $j < i \leq n-1$. Then

$$\sum_{i=0}^{j}(10S)^i x_i = \sum_{i=0}^{j}(10S)^i y_i.$$

Moving the x_j and y_j to one side as above, we see that $x_j = y_j$.

Next we prove that B contains no 3-term arithmetic progression. Suppose that $\phi(x), \phi(y), \phi(z)$ form an arithmetic progression, where $x, y, z \in A$. Using a similar argument to the proof of injectivity, we will show that x, y, z form an arithmetic progression too. Since there are no 3-term arithmetic progressions in A, we can conclude that there are no 3-term arithmetic progressions in B. Now $\phi(x), \phi(y), \phi(z)$ form a 3-term arithmetic progression if and only if

$$\phi(y) - \phi(x) = \phi(z) - \phi(y),$$

if and only if

$$\sum_{i=0}^{n-1}(10S)^i(y_i - x_i) = \sum_{i=0}^{n-1}(10S)^i(z_i - y_i).$$

Noting that $|y_i - x_i|, |z_i - y_i| \leq 2S$, and using the argument above, we see that for every i,

$$y_i - x_i = z_i - y_i.$$

But this shows that x, y, z form a 3-term arithmetic progression. This finishes the proof of Theorem 3.11.

The construction of Behrend's example involves polynomials, although it is not clear to me whether they play a crucial role. A key step in the argument was to find a strictly convex surface $\Sigma \subset \mathbb{R}^n$ which contains many points of the grid $[-S, ..., S]^n$. This leads to the question: among all strictly convex hypersurfaces $\Sigma \subset (-S, S)^n$, what is the maximum possible number of lattice points in Σ? In the construction of the Behrend example, we used a sphere. In order to check that some sphere contains many lattice points, we examined the polynomial equation $\sum_i x_i^2 = M$. Are there other strictly convex surfaces with more lattice points than a sphere? Without mentioning polynomials, can one find an example which is comparable to a sphere?

3.4. What is special about polynomials?

Now that we tried for a little while to work on the finite field Kakeya problem and the joints problem without polynomials, we return to the polynomial proofs, and we examine what properties of polynomials made them well-suited to these problems. I would like to highlight two facts which play a crucial role in the argument.

- $\operatorname{Dim} \operatorname{Poly}_D(\mathbb{F}^n) \sim D^n$.
- If $P \in \operatorname{Poly}_D(\mathbb{F}^n)$ vanishes at more than D points of a line $l \subset \mathbb{F}^n$, then P vanishes on the whole line l.

The first key fact says that there are lots of polynomials. The second key fact says that polynomials behave rather rigidly on lines. When we pick a polynomial $P \in \operatorname{Poly}_D(\mathbb{F}^n)$, we have $\sim D^n$ degrees of freedom at our disposal, and this gives us a lot of flexibility. But then, when we consider P restricted to a line, it behaves surprisingly rigidly, with only $\sim D$ degrees of freedom. These facts show that polynomials have a special relationship with lines. The gap between D^n and D gives us a kind of leverage which powers the proofs in Chapter 2.

Here is a more precise statement of the fact that a polynomial restricted to a line has few degrees of freedom. If $W \subset \operatorname{Fcn}(\mathbb{F}^n, \mathbb{F})$ is a vector space of functions on $\mathbb{F}^n$, and $X \subset \mathbb{F}^n$, then we define the dimension of W restricted to X as the rank of the evaluation map E_X from W to $\operatorname{Fcn}(X, \mathbb{F})$. In symbols,

$$\operatorname{Dim} W|_X := \operatorname{Rank} \left[E_X : W \to \operatorname{Fcn}(X, \mathbb{F}) \right].$$

For any line $l \subset \mathbb{R}^n$, $\operatorname{Dim} \operatorname{Poly}_D(\mathbb{F}^n)|_l = D + 1$. By contrast, $\operatorname{Dim} \operatorname{Poly}_D(\mathbb{F}^n) \sim D^n$.

Most function spaces do not have any such gap. For example, let us consider the trigonometric polynomials of degree $\leq D$ on $\mathbb{R}^n$. Recall that a trigonometric polynomial of degree $\leq D$ is a function $f : \mathbb{R}^n \to \mathbb{C}$ of the following form:

$$f(x) = \sum_{\omega \in \mathbb{Z}^n, |\omega_j| \leq D \text{ for all } j} a(\omega) e^{2\pi i \omega \cdot x}.$$

Here the coefficients $a(\omega)$ are complex numbers. We let $\operatorname{Trig}_D(\mathbb{R}^n)$ denote the vector space of trigonometric polynomials on $\mathbb{R}^n$ of degree at most D. The dimension of $\operatorname{Trig}_D(\mathbb{R}^n)$ is $\sim D^n$, like the dimension of $\operatorname{Poly}_D(\mathbb{R}^n)$. Moreover, if l is an axis-parallel line, then the restriction of f to l is a trigonometric polynomial in one variable of degree at most D. So if l is an axis-parallel line, then $\operatorname{Dim} \operatorname{Trig}_D(\mathbb{R}^n)|_l \sim D$. But for a generic line $l \subset \mathbb{R}^n$, $\operatorname{Dim} \operatorname{Trig}_D(\mathbb{R}^n)|_l = \operatorname{Dim} \operatorname{Trig}_D(\mathbb{R}^n)$. For generic lines, trigonometric polynomials don't offer any "leverage" in the sense above.

If this "leverage" is the key to the polynomial method proofs in Chapter 2, then it is natural to try to find the function space with the most leverage. We formalize this question as follows.

QUESTION 3.13. Fix a field $\mathbb{F}$ and a dimension n. Suppose that $W \subset \operatorname{Fcn}(\mathbb{F}^n, \mathbb{F})$ is a vector space of functions, and suppose that for every line $l \subset \mathbb{F}^n$,

$$\operatorname{Dim} W|_l \leq D + 1.$$

What is the maximum possible dimension of W?

I don't know anything about this question. But here is a closely related question which is better understood. We say that W obeys the degree D vanishing lemma if, for any $f \in W$, if $f = 0$ at $D+1$ points of a line, then $f = 0$ at every point on the line.

Question 3.14. What is the maximum possible dimension of a vector space of functions $W \subset \mathrm{Fcn}(\mathbb{F}^n, \mathbb{F})$ which obeys the degree D vanishing lemma?

Exercise 3.3. If $|\mathbb{F}| \geq D+1$, and if $W \subset \mathrm{Fcn}(\mathbb{F}^n, \mathbb{F})$ obeys the degree D vanishing lemma, prove that $\mathrm{Dim}\, W \leq (D+1)^n$.

In the lecture notes of a course on the polynomial method, I conjectured that $Poly_D(\mathbb{F}^n)$ has the largest possible dimension among all spaces obeying the degree D vanishing lemma. This turns out to be false, at least for some fields $\mathbb{F}$. In [**LuSu**], Luo and Sudan constructed larger vector spaces of functions that still obey the degree D vanishing lemma. Using these function spaces instead of polynomials, they were able to give sharper constants in the finite field Nikodym theorem.

Polynomials are a special class of functions. They play an important role in many areas of mathematics outside of algebra, and each area has its own perspective about what makes them special. Over the course of the book, we will explore a couple of other areas where polynomial play a special role, because ideas in those areas are related to the polynomial method in combinatorics. In Chapter 4, we will talk about error-correcting codes in computer science. In this field, people talk about the 'resiliency' of polynomials: you can distort or damage a polynomial and there is enough information left to recover the original polynomial. This makes them important tools in error-correcting codes. In Chapter 14, we will talk about zero sets of polynomials in differential geometry. In this field, people talk about the 'efficiency' of polynomials: the zero sets of polynomials have minimal size or minimal complexity in several different ways. Both the 'resiliency' of polynomials and the 'efficiency' of polynomials are connected with the two special features we have discussed here: the dimension of $\mathrm{Poly}_D(\mathbb{F}^n)$ and the vanishing lemma.

3.5. An example involving polynomials

Since polynomials play an important role in the proofs, it's reasonable to ask if there are interesting examples of configurations of lines that are based on polynomials. The Hermitian variety is one important example. In this section, we describe an even simpler example that will play an important role in the book. This example is a configuration of lines in $\mathbb{R}^3$ based on a degree 2 algebraic surface.

To motivate this example, we begin with some naive questions about the intersection patterns of lines in $\mathbb{R}^3$. Suppose we have L lines in $\mathbb{R}^3$. How many intersection points can there be? There are at most $\binom{L}{2}$ intersection points, and this can be achieved by putting all the lines in a plane.

What if we don't allow ourselves to put all the lines in a plane? Suppose we have L lines in $\mathbb{R}^3$ with ≤ 10 lines in any plane. How many intersection points can there be? Remarkably, there can still be $\sim L^2$.

Let S be the degree 2 algebraic surface defined by the equation $z = xy$. The surface S contains many lines. For each y_0, there is a 'horizontal line' $h(y_0) \subset S$ parametrized by $\gamma(t) = (t, y_0, y_0 t)$. And for each x_0, there is a 'vertical line' $v(x_0) \subset S$ parametrized by $\gamma(t) = (x_0, t, x_0 t)$. Any horizontal line intersects any vertical line: $h(y_0)$ intersects $v(x_0)$ at $(x_0, y_0, x_0 y_0)$. Moreover, all these intersection points

are distinct. Taking $L/2$ horizontal lines and $L/2$ vertical lines gives $L^2/4$ distinct intersection points. On the other hand, any plane intersects the surface S, and so any plane contains at most 2 of our lines.

The surface S is an example of a regulus. We will study reguli in Section 8.4. Reguli played a crucial role in the first work on the joints problem, which we will describe there.

This is an important example in combinatorial problems about intersecting lines. It shows that interesting examples don't come only from subspaces and objects of linear algebra - they also come from low degree algebraic surfaces. This example helps motivate using polynomials to study the combinatorics of lines in $\mathbb{R}^3$. If examples that we are worried about can come from polynomials, then we may hope to enlist the aid of polynomials either to find such examples or to rule them out.

3.6. Combinatorial structure and algebraic structure

We see that a degree 2 polynomial leads to an interesting configuration of lines. Next we may wonder if all the interesting configurations come from polynomials.

Continuing our naive questions, what if we forbid the lines to cluster in planes or degree 2 surfaces? More formally, we have the following question:

QUESTION 3.15. Suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with ≤ 10 lines in any plane or degree 2 algebraic surface. What is the maximum possible number of intersection points of $\mathfrak{L}$?

This time, the answer is much less than L^2. Getting optimal bounds is an open question that looks important to me. The best currently known upper bound is $\sim L^{3/2}$. This upper bound was proven in [**GK2**], and we will prove it in Chapter 13. The only examples I know have far fewer intersection points.

We can get some perspective on this problem by counting parameters. The set of lines in $\mathbb{R}^3$ is a 4-dimensional manifold. So choosing L lines gives us $4L$ parameters to play with. If we want one particular line to intersect another, that gives us one equation that our parameters have to satisfy. Just counting parameters, one might guess that it's not hard to find examples with $4L$ intersections. On the other hand, one might guess that examples with far more than $4L$ intersection points should come from some special structure. Question 3.15 asks whether that special structure needs to be a plane or a degree 2 surface.

A little more generally, we can ask the following question:

QUESTION 3.16. Let $\mathfrak{L}$ be a set of L lines in $\mathbb{R}^3$ with at most B lines in any algebraic surface of degree at most D, what is the maximum possible number of intersection points of $\mathfrak{L}$?

This is one of the central problems of the book. We will come back to it three times with three different approaches: we will use reguli in Chapter 8, we will use polynomial partitioning in Chapter 10, and we will use ruled surface theory in Chapter 13. As we develop more tools, our estimates will get stronger. In the regime $B \geq L^{1/2}$, we will be able to prove sharp upper bounds - cf. Theorem 8.3. On the other hand, we remain far from a complete answer, and this problem marks an important boundary in our understanding of the subject.

The theme of this section is the connection between combinatorial structure and algebraic structure. We have seen some examples with a lot of combinatorial

structure which are built using polynomials. Does every example with a lot of polynomial structure come from some algebraic structure? Questions 3.15 and 3.16 are precise questions that get at this issue. Investigating this issue is one of the main goals of the book.

https://doi.org/10.1090//ulect/064/04

CHAPTER 4

The polynomial method in error-correcting codes

The proofs of the finite field Kakeya theorem and the joints theorem drew on ideas from computer science. Polynomials over finite fields have been studied intensively by computer scientists. Polynomials over finite fields are also classical mathematical objects, and they've been studied intensively by mathematicians for many years. But problems in computer science have suggested different types of questions and led to new perspectives on polynomials.

The two main ingredients in the proofs of finite field Nikodym and joints are the parameter counting lemma and the vanishing lemma. This team of ingredients appeared together earlier in the theory of error-correcting codes. In this chapter, we present a few interesting results from error-correcting codes, illustrating these techniques.

4.1. The Berlekamp-Welch algorithm

Let $\mathbb{F}_q$ be the finite field with q elements, and let $\mathrm{Poly}_D(\mathbb{F}_q)$ be the vector space of all polynomials in one variable with degree $\leq D$. Because of the vanishing lemma, any two different polynomials in $\mathrm{Poly}_D(\mathbb{F}_q)$ can only agree at $\leq D$ values of $x \in \mathbb{F}_q$. If D is much less than q, then any two polynomials in $\mathrm{Poly}_D(\mathbb{F}_q)$ look very different from each other. This makes them interesting tools for error-correcting codes.

Here is a typical situation from coding theory. Suppose Q is a polynomial over $\mathbb{F}_q$ with degree $\leq q/100$. We want to transmit or save Q, but the data gets corrupted and instead we end up with a function $F : \mathbb{F}_q \to \mathbb{F}_q$. Suppose we know that F agrees with Q for a certain fraction of $x \in \mathbb{F}_q$. Is it possible to recover Q from F? Is it possible to do it efficiently?

An interesting case is when F agrees with Q a little more than half the time. Let's suppose that $F(x) = Q(x)$ for at least $(51/100)q$ values of x. In this case, it follows immediately from the vanishing lemma that we can recover Q from F in theory.

LEMMA 4.1. Let $F : \mathbb{F}_q \to \mathbb{F}_q$ be any function. Then there is at most one polynomial $Q \in \mathrm{Poly}_{q/100}(\mathbb{F}_q)$ which agrees with F for $\geq (51/100)q$ values of x.

PROOF. Suppose that $Q_1, Q_2 \in \mathrm{Poly}_{q/100}(\mathbb{F}_q)$ both agree with F for at least $(51/100)q$ values of x. But then $Q_1(x) = Q_2(x)$ for $\geq (2/100)q$ values of x. Now $Q_1 - Q_2$ vanishes at $\geq (2/100)q$ points but has degree $\leq (1/100)q$. By the vanishing lemma, $Q_1 - Q_2$ is the zero polynomial. □

The function F does contain enough information to recover Q in theory. But there is a deeper question: can we recover Q from F in an efficient way? We could

find Q by trying all the polynomials in $\mathrm{Poly}_{q/100}(\mathbb{F}_q)$. But the number of polynomials in this set grows more than exponentially fast in q, and so the running time of this naive algorithm is more than exponential in q. In the mid-80's, Berlekamp and Welch [**BW**] gave a much more efficient algorithm to recover Q from F. The running time of their algorithm is polynomial in q, and the algorithm is fast enough to be useful in practice. Their solution combines the parameter counting idea and the vanishing lemma in an elegant way.

THEOREM 4.2. (Berlekamp-Welch, [**BW**], 1986) Suppose that $Q(x)$ is a polynomial over $\mathbb{F}_q$ with degree $< q/100$. Suppose that $F : \mathbb{F}_q \to \mathbb{F}_q$, and that $F(x) = Q(x)$ for $\geq (51/100)q$ values of x. Then there is a polynomial time algorithm that recovers Q from F.

This error-correcting code is called a Reed-Solomon code. The message we want to encode is a sequence of $D+1$ elements of $\mathbb{F}_q$: $a_0, ..., a_D$. Using this sequence, we build a polynomial $Q(x) = \sum_{j=0}^{D} a_j x^j$. Then we record the values of Q at each point of $\mathbb{F}_q$. Even if 49 % of the recorded data is corrupted by a clever adversary, we can still recover the original message in polynomial time (provided that $D < q/100$).

The Berlekamp-Welch algorithm is based on studying the graph of F: the set $\{(x, y) \in \mathbb{F}_q^2 | F(x) = y\}$. This graph looks like a cloud of points. Inside the cloud of points a certain algebraic structure is hidden: most of the points lie on the graph of Q. How can we search out this algebraic structure hidden in the cloud of points?

The main idea of the algorithm is to find the lowest degree non-zero polynomial $P(x, y)$ that vanishes on the graph of F. On the one-hand, we can find an optimal P with an efficient algorithm. On the other hand, this optimal P uncovers the hidden algebraic structure in the cloud of points: looking at the zero set of P, the graph of Q jumps off the page.

In our algorithm, we will treat the degree of P in x and the degree of P in y differently. This makes sense if we look back at the problem we're trying to solve. We're hoping to find the graph of Q, which is defined by the equation $y - Q(x) = 0$. This defining equation has degree 1 in y and high degree in x. In order to adapt to the problem, it turns out to be a good idea to use polynomials $P(x, y)$ of degree 1 in y and high degree in x. We let $\mathrm{Poly}_{D,E}(\mathbb{F}_q^2)$ be the vector space of polynomials $P(x, y)$ with degree $\leq D$ in x and $\leq E$ in y.

The first step of the algorithm is to find a polynomial $P(x, y)$ which vanishes on the graph of F, where P has degree 1 in y and the smallest possible degree in x. We can do this by the following Proposition:

PROPOSITION 4.3. There in a polynomial time algorithm that does the following. Given any set $S \subset \mathbb{F}_q^2$, the algorithm finds a non-zero polynomial $P(x, y) \in \mathrm{Poly}_{D,1}(\mathbb{F}_q^2)$ which vanishes on S and where the degree D is as small as possible. The degree D will always obey $D \leq |S|/2$.

PROOF. This problem boils down to linear algebra.

For a given D, we consider the restriction map $R_S : \mathrm{Poly}_{D,1}(\mathbb{F}_q^2) \to \mathrm{Fcn}(S, \mathbb{F}_q)$, which restricts a polynomial $P \in \mathrm{Poly}_{D,1}(\mathbb{F}_q^2)$ to the set $S \subset \mathbb{F}_q^2$. The map R_S is a linear map, and we want to check whether it has a non-trivial kernel. We can write R_S as a matrix after we choose a basis for the domain and for the range. For the domain, a natural basis is given by the monomials $x^a y^b$, with $0 \leq a \leq D$ and $0 \leq b \leq 1$. For the range, a natural basis is given by delta functions at the points of

S. If $S = \{(x_1, y_1), (x_2, y_2), ...\}$ then the basis elements of S are $\{\delta_{(x_j,y_j)}\}_{j=1...,|S|}$. In this basis, the matrix entry corresponding to the row (x_j, y_j) and the column (a, b) is just $x_j^a y_j^b$. With this explicit matrix, we can compute a basis for the kernel of R_S by using Gaussian elimination. The running time is polynomial in the dimensions of the matrix. For an $N \times N$ matrix, Gaussian elimination takes time $\sim N^3$, because it involves $\sim N^2$ row and column operations, and each such operation involves $\sim N$ computations.

We perform this calculation for $D = 0$, then for $D = 1$, etc. The dimension of $\text{Poly}_{D,1}(\mathbb{F}_q^2)$ is $2D + 2$. If $2D + 2 > |S|$, then the kernel of R_S is guaranteed to be non-trivial. Therefore, we will find a non-trivial kernel for some D in the range $0 \leq D \leq |S|/2$. We note the lowest value of D so that R_S has a non-trivial kernel, and we let $P(x, y) \in \text{Poly}_{D,1}(\mathbb{F}_q^2)$ be a non-zero element of this kernel.

□

We apply Proposition 4.3 with the set S being the graph of F. The number of points in the graph is q, and so we have $D \leq q/2$. We can write $P(x, y) = P_0(x) + yP_1(x)$ with $\text{Deg}\, P_0, \text{Deg}\, P_1 \leq q/2$. Now the key point in the proof of Theorem 4.2 is that $P(x, y)$ vanishes on the graph of Q. This follows in a few simple steps.

1. We know $P = 0$ on the graph of F. In other words, $P(x, F(x)) = 0$ for all x.

2. But we know that F usually agrees with Q. So $P(x, Q(x)) = 0$ for at least $(51/100)q$ values of x.

3. But $P(x, Q(x)) = P_0(x) + Q(x)P_1(x)$ is a polynomial in x of degree $\leq \text{Deg}\, Q + \max(\text{Deg}\, P_0, \text{Deg}\, P_1) < q/100 + q/2 = (51/100)q$.

4. By the vanishing lemma, $P(x, Q(x))$ is the zero polynomial.

We have proven that $P(x, Q(x))$ is identically zero, and so P vanishes on the graph of Q. Moreover, since $0 = P(x, Q(x)) = P_0(x) + Q(x)P_1(x)$, we see that $Q(x)P_1(x) = -P_0(x)$. We know P_0 and P_1, and now we can recover Q by doing polynomial division. This is the Berlekamp-Welch algorithm.

There is a more visual way of explaining how to recover Q, which makes the graph of Q jump off the page. We let the set of errors be $E := \{x \in \mathbb{F}_q | F(x) \neq Q(x)\}$. With a little more work, we will prove the following claim:

CLAIM 4.4. $P(x, y) = c\,[y - Q(x)] \prod_{e \in E}(x - e)$, for some non-zero constant $c \in \mathbb{F}$.

This claim implies that the zero set of our polynomial P is the union of the graph of Q and a vertical line $x = e$ at each error $e \in E$. Looking at the zero set of P, the set of errors is immediately visible, together with a large chunk of the graph of Q. From this large chunk of the graph of Q, we can quickly recover Q itself.

The proof of the claim uses a divisibility lemma:

LEMMA 4.5. If $P(x, y)$ is a polynomial of two variables, and $Q(x)$ is a polynomial in one variable, and $P(x, Q(x))$ is the zero polynomial, then $P(x, y) = (y - Q(x))P_1(x, y)$ for some polynomial P_1.

Let us prove Claim 4.4 using the divisibility lemma. We saw above that $P(x, Q(x))$ is the zero polynomial – by Lemma 4.5, $P(x, y) = (y - Q(x))P_1(x)$. Now if $e \in E$, then $Q(e) \neq F(e)$. We know that P vanishes on the graph of F, and so

$$0 = P(e, F(e)) = (F(e) - Q(e))P_1(e),$$

and so $P_1(e) = 0$. Therefore, P_1 is divisible by $x - e$ for each $e \in E$, and we see that

$$P(x, y) = (y - Q(x)) \prod_{e \in E} (x - e) P_2(x).$$

Since P has minimal degree, P_2 must be a non-zero constant. This finishes the proof of Claim 4.4.

We end this section by giving the proof of the divisibility lemma, Lemma 4.5. This argument is similar to the proof of the vanishing lemma, Lemma 2.8. That proof was based on Lemma 2.6, which says the following:

LEMMA. If $P(y) \in \mathrm{Poly}_D(\mathbb{F})$ is a polynomial in one variable and $y_1 \in FF$, then we can write P in the form

$$P(y) = (y - y_1) P_1(y) + r,$$

where $P_1(y) \in \mathrm{Poly}_{D-1}(\mathbb{F})$ and $r \in \mathbb{F}$.

We now give a version of this lemma for polynomials of two variables.

LEMMA 4.6. If $P(x, y) \in \mathrm{Poly}(\mathbb{F}^2)$ has $\mathrm{Deg}_y P \leq D$, and $Q(x) \in \mathrm{Poly}(\mathbb{F})$, then we can write P in the form

$$P(x, y) = (y - Q(x)) P_1(x, y) + R(x),$$

where $\mathrm{Deg}_y P_1(x, y) \leq D - 1$ and $R \in \mathrm{Poly}(\mathbb{F})$.

PROOF. We do the proof by induction on D. If $D = 0$, then $P(x, y) = R(x)$ and the conclusion is clear.

Suppose $P(x, y) = \sum_{j=0}^{D} a_j(x) y^j$, where $a_j \in \mathrm{Poly}(\mathbb{F})$. Let $\tilde{P}(x, y) = P(x, y) - (y - Q(x))(a_D(x) y^{D-1})$. The y^D term of $\tilde{P}(x, y)$ vanishes, and so $\mathrm{Deg}_y \tilde{P} \leq (D-1)$. By induction on D, we can write

$$P(x, y) - (y - Q(x))(a_D(x) y^{D-1}) = \tilde{P}(x, y) = (y - Q(x)) \tilde{P}_1(x, y) + R(x),$$

where $\mathrm{Deg}_y \tilde{P}_1 \leq D - 1$ and $R \in \mathrm{Poly}(\mathbb{F})$. Therefore, we see

$$P(x, y) = (y - Q(x))(a_D(x) y^{D-1} + \tilde{P}_1(x, y)) + R(x).$$

□

This lemma quickly implies Lemma 4.5:

PROOF OF LEMMA 4.5. Suppose that $P(x, Q(x))$ is the zero polynomial. By the previous lemma, we can write P in the form

$$P(x, y) = (y - Q(x)) P_1(x, y) + R(x).$$

Plugging in $y = Q(x)$, we see that $P(x, Q(x)) = R(x)$. Therefore R is the zero polynomial, and $P(x, y) = (y - Q(x)) P_1(x, y)$. □

4.2. Correcting polynomials from overwhelmingly corrupted data

In the Berlekamp-Welch algorithm, we considered corrupted data F which was correct a little more than half the time. If F is correct only half the time, then it's impossible to recover the polynomial Q even in theory. For example, start with two low degree polynomials Q_1 and Q_2, and arrange for F to agree with Q_1 half the time and with Q_2 half the time. There is no way to tell if the original polynomial was Q_1 or Q_2. Following this observation, it may seem that data F which is correct

only 1 % of the time would not be very useful. Surprisingly, it turns out that a great deal of information can be recovered from such data. In the mid 90's, Sudan generalized the algorithm of Berlekamp-Welch to deal with highly corrupted data. For example, he proved the following result.

THEOREM 4.7. (Sudan, 1997) Suppose that $\mathbb{F}$ is a field with q elements, and that $F : \mathbb{F} \to \mathbb{F}$ is any function. There is an efficient algorithm that lists all the polynomials of degree $< (1/200)q^{1/2}$ that agree with F for at least $q/100$ values of x.

We have the tools to follow most of the steps of Sudan's argument. We again consider the graph of F in $\mathbb{F}^2$. We find a low-degree polynomial $P(x, y)$ that vanishes on the graph. By the same argument as in the proof of Proposition 4.3, we can efficiently find a non-zero polynomial $P(x, y) \in \mathrm{Poly}_D(\mathbb{F}_q^2)$ that vanishes on the graph of Q, and where the degree D is as small as possible. Since the graph of Q has q elements, Lemma 2.4 tells us that there is a non-zero polynomial vanishing on the graph of Q with degree at most $2q^{1/2}$. In particular, the degree of P is $\leq 2q^{1/2}$.

Suppose that Q has degree $< (1/200)q^{1/2}$, and that $Q(x) = F(x)$ for at least $q/100$ values of x. We claim that $P(x, Q(x))$ is the zero polynomial. This follows for the same reason as above. We know that $P(x, F(x))$ is zero for every x. So $P(x, Q(x))$ has at least $q/100$ zeroes. But $P(x, Q(x))$ is a polynomial of degree at most $(\mathrm{Deg}\, P)(\mathrm{Deg}\, Q) < 2q^{1/2}(1/200)q^{1/2} = q/100$. Therefore $P(x, Q(x))$ is the zero polynomial. By the divisibility lemma, Lemma 4.5, we see that $y - Q(x)$ divides $P(x, y)$.

We have efficiently constructed $P(x, y)$, and we know that for every low degree Q which agrees with F at $\geq q/100$ places, $y - Q(x)$ divides $P(x, y)$.

There is a polynomial time algorithm that factors $P(x, y)$ into irreducible factors. This step is not at all obvious, and it requires different ideas, cf. [**Ka**]. With this algorithm, we can find all factors of $P(x, y)$ of the form $y - Q(x)$. The number of such factors is at most $\mathrm{Deg}\, P \leq 2q^{1/2}$. Finally we check all of the $Q's$ that arise from the factorization of $P(x, y)$ and we see which ones agree with F for at least $q/100$ values of x. This list is the output of our algorithm.

4.3. Locally decodable codes

The Reed-Solomon code, the error-correcting code we considered in the first section, is quite robust: even if 49 % of the recorded data is corrupted by an adversary, we can still recover the original message. On the other hand, this code is a little unwieldy in the following sense: whenever we want to read off a single letter of the original message from the recorded data, we need to decode the entire message - a process which will certainly take at least as long as reading the entire message. In a locally decodable code, any single letter of the message can be quickly decoded, while looking at only a small fraction of the recorded data, in time far less than it would take to read through the original message.

It would be interesting to find a locally decodable code which is still as robust as the Reed-Solomon code: even after 49 % of the recorded data is corrupted, we would like to be able to quickly read off any desired letter of the original message by looking at only a small fraction of the recorded data. For this to work, there

need to be many different ways to reconstruct each letter of the message, and these different ways need to draw on different parts of the recorded data.

The Reed-Muller code is an interesting code that achieves this goal. It is based on polynomials on $\mathbb{F}_q^n$ instead of polynomials on $\mathbb{F}_q$. Here is the definition of the code.

For any $D < q$ and any $n \geq 1$, we will construct a code. The original message is a list of $(D+1)^n$ elements of $\mathbb{F}_q$, which we think of as a function

$$g : \{0, ..., D\}^n \to \mathbb{F}_q.$$

Each such function extends to a unique polynomial $P : \mathbb{F}_q^n \to \mathbb{F}_q$ with degree at most D in each variable. We will prove this as a lemma in a moment. The recorded data of the Reed-Muller code is the polynomial P.

LEMMA 4.8. If $D < q$, then for any function $g : \{0, ..., D\}^n \to \mathbb{F}_q$, there is a unique polynomial $P : \mathbb{F}_q^n \to \mathbb{F}_q$ so that $P = g$ on $\{0, ..., D\}^n$ and $\mathrm{Deg}_{x_i} P \leq D$ for each $i = 1, ..., n$.

PROOF. The proof is by induction on n with base case $n = 1$.

If $n = 1$, the proof follows by linear algebra and van der Monde determinants. It is closely related to the proof of the vanishing lemma, Lemma 2.8. Let $S := \{0, ..., D\} \subset \mathbb{F}_q$. We want to show that the evaluation map $E_S : \mathrm{Poly}_D(\mathbb{F}_q) \to \mathrm{Fcn}(S, \mathbb{F}_q)$ is an isomorphism. We note that E_S is a linear map and that the domain and target both have dimension $D + 1$. We will express the linear map as a matrix and then check that its determinant is not zero. We write a polynomial $Q \in \mathrm{Poly}_D(\mathbb{F}_q)$ in the standard form $Q(x) = \sum_{j=0}^{D} a_j x^j$. For each $i \in S$, we have $E_S(Q)(i) = Q(i) = \sum_{j=0}^{D} i^j a_j$. Therefore, the restriction map E_S is given by a $(D+1) \times (D+1)$ matrix M with coefficients $M_{ij} = i^j$, for $0 \leq i, j \leq D$. The determinant of this matrix is given by the van der Monde determinant formula:

$$\det M = \prod_{0 \leq i_1 < i_2 \leq D} (i_1 - i_2) \neq 0.$$

Therefore E_S is an isomorphism, and this shows that the Lemma is true in the base case $n = 1$.

We now turn to the higher-dimensional case. Using the case $n = 1$, we see that for any $(x_1, ..., x_{n-1}) \in \{0, ..., D\}^{n-1}$, there is a unique choice of coefficients $a_0(x_1, ..., x_{n-1}), ..., a_D(x_1, .., .x_{n-1}) \in \mathbb{F}_q$ so that

$$g(x_1, ..., x_n) = \sum_{j=0}^{D} a_j(x_1, ..., x_{n-1}) x_n^j \text{ for all } x_n = 0, ..., D.$$

Note that $a_j : \{0, ..., D\}^{n-1} \to \mathbb{F}_q$. By induction on n, there is a unique polynomial $P_j : \mathbb{F}_q^{n-1} \to \mathbb{F}^q$ so that $P_j = a_j$ on $\{0, ..., D\}^{n-1}$ and $\mathrm{Deg}_{x_i} P_j \leq D$ for each $1 \leq i \leq n-1$. Now for all $(x_1, ..., x_n) \in \{0, ..., D\}^n$, we have $g(x_1, ..., x_n) = \sum_{j=0}^{D} P_j(x_1, .., x_{n-1}) x_n^j = P(x_1, ..., x_n)$. We also see that $\mathrm{Deg}_{x_i} P \leq D$ for each $i = 1, ..., n$. So the polynomial P satisfies our conditions.

Finally, to check that P is unique, suppose that Q agrees with G on $\{0, ..., D\}^n$ and has $\mathrm{Deg}_{x_i} Q \leq D$ for all $1 \leq i \leq n$. We can write Q in the form

$$Q(x) = \sum_{j=0}^{D} Q_j(x_1, ..., x_{n-1}) x_n^j,$$

where Q_j is a polynomial of degree at most D in each variable. By the discussion above, we must have $Q_j(x_1, ..., x_{n-1}) = a_j(x_1, ..., x_{n-1})$ for each $(x_1, ..., x_{n-1}) \in \{0, ..., D\}^{n-1}$, and so $Q_j = P_j$, and so $Q = P$.

□

The proof shows how to find P from g using linear algebra, and this can be done in polynomial time. Next we describe how to efficiently recover $g(x)$ from a corrupted version of P, following [**Su**]. First, we describe how to recover from 24 % error.

Suppose that $g : \{0, ..., D\}^n \to \mathbb{F}^q$ is a function, and that $P : \mathbb{F}_q^n \to \mathbb{F}_q$ extends g with $\mathrm{Deg}_{x_i} P \leq D$ for each i. Suppose that $F(x) = P(x)$ for at least $(76/100)q^n$ elements of $\mathbb{F}_q^n$. Suppose that we are given $x \in \{0, ..., D\}^n$, and we wish to recover $g(x)$ from the function F, in a quick way that reads F at far less than q^n points.

We randomly pick a line ℓ through the point x. We let F_ℓ be the restriction of F to ℓ. We apply the Berlekamp-Welch algorithm to F_ℓ. Note that the degree of P is at most nD. We suppose that $nD < q/100$. Now as long as $F_\ell(y) = P_\ell(y)$ for at least $(51/100)q$ of the points $y \in \ell$, the Berlekamp-Welch algorithm will recover P_ℓ from F_ℓ. In this case, we call ℓ a good line. If ℓ is a good line, the Berlekamp-Welch algorithm will tell us $P(x) = g(x)$. If ℓ is not a good line, then the Berlekamp-Welch algorithm may either output an incorrect polynomial $Q_\ell \neq P_\ell$ or it may terminate with a message that F_ℓ was not actually close to a low degree polynomial.

As long as q is sufficiently large, we claim that at least 51% of the lines ℓ through x are good. Checking the claim is just a simple computation. If the claim is false, there are at least $(49/100)q^{n-1}$ lines ℓ through x that contain at least $(49/100)q - 1$ points $y \in \ell \setminus \{x\}$ where $F(y) \neq P(y)$. The total number of such bad points is at least $(.49)^2 q^n - O(q^{n-1}) = (.2401)q^n - O(q^{n-1})$. If q is sufficiently large, this is $> (.24)q^n$.

This leads to a randomized algorithm that finds $g(x)$ with high probability and only reads $F(x)$ in $\lesssim q$ places. We pick A random lines ℓ through x. For each random line, we use the Berlekamp-Welch algorithm to try to decode $g(x)$. With probability at least 51%, a line ℓ is good and delivers the correct value of $g(x)$. Our algorithm outputs the most popular guess for $g(x)$ among the A lines. The probability of error decays exponentially in A, and we can arrange 99.9 % certainty of correctness by taking a moderately large constant A. The algorithm only reads F at Aq points.

The running time of the algorithm depends on the running time of Berlekamp-Welch. With the simple description of Berlekamp-Welch from this chapter, the running time would be $\sim q^4$. If we take n much larger than 4, then the running time of the algorithm to decode $g(x)$ can be much shorter than D^n - the time it would take to read the original message.

By being a little trickier, we can recover $g(x)$ with a similar algorithm even if we have 49 % error. We will describe the main modification and leave the proof as an exercise. To motivate what we do, let us explain why our first algorithm can fail with 49 % error. Our adversary picks a point $x \in \{0, ..., D\}^n$ and designs F to confuse us when we try to recover $g(x)$. The adversary is allowed to introduce errors at 49 % of the points $x \in \mathbb{F}_q^n$. The adversary picks a polynomial $\tilde{P} : \mathbb{F}_q^n \to \mathbb{F}^q$ with degree at most D in each variable. Now for 95 % of the lines ℓ through x, the adversary arranges that $F(y) = \tilde{P}(y)$ for at least $(51/100)q$ of the points $y \in \ell$. For the remaining 5 % of the lines ℓ through x, the adversary can only introduce very

few errors. When we pick a random line ℓ through x, run Berlekamp-Welch, and guess the value of $g(x)$, 95 % of the lines give the guess $\tilde{P}(x) \neq g(x)$, and only 5 % of the lines give to the correct guess $P(x) = g(x)$.

However, there is a clue that could help us realize that the 5 % of lines that give the guess $g(x)$ are more reliable than the 95 % of lines that give the wrong guess. The Berlekamp-Welch algorithm outputs a polynomial Q_ℓ defined on the line ℓ. For each of bad lines, Q_ℓ agrees with F_ℓ at only $(51/100)q$ points $y \in \ell$. But for each good line, Q_ℓ agrees with F_ℓ at almost every point $y \in \ell$. We can make a better algorithm by weighting the vote of a line more heavily if Q_ℓ agrees with F_ℓ at a large fraction of the points $y \in \ell$.

Here is a modified algorithm to recover $g(x)$ from the function F. We pick A random lines ℓ through x. For each of these lines, we run the Berlekamp-Welch algorithm. For a given line ℓ, if the algorithm outputs a polynomial Q_ℓ, then we proceed as follows. We count the number of points $y \in \ell$ where $F_\ell(y) = Q_\ell(y)$. We let $p(\ell)$ be the fraction of points $y \in \ell$ where $F_\ell(y) = Q_\ell(y)$, and we define a weight $w(\ell) = 2p(\ell) - 1$. Then we say that the line ℓ votes for the value $Q_\ell(x)$ with a vote of weight $w(\ell)$. (If the algorithm terminates with a message that F_ℓ was not actually close to a low degree polynomial, then the line ℓ does not vote for any value.) We output the winner of the election: the value that receives votes of the greatest total weight.

EXERCISE 4.1. Suppose that $g : \{0, ..., D\}^n \to \mathbb{F}^q$ is a function, and that $P : \mathbb{F}_q^n \to \mathbb{F}_q$ extends g with $\mathrm{Deg}_{x_i} P \leq D$ for each i. Suppose that q is sufficiently large, that $nD < q/100$. Suppose that $F(x) = P(x)$ for at least $(51/100)q^n$ elements of $\mathbb{F}_q^n$. For any $x \in \{0, ..., D\}^n$, prove that the algorithm described in the last paragraph will recover $g(x)$ with high probability.

4.4. Error-correcting codes and finite-field Nikodym

The proofs of the finite field Nikodym conjecture and the finite field Kakeya conjecture were partly inspired by ideas from error-correcting codes. After discussing some of these ideas, let us revisit the proof of the finite field Nikodym conjecture from the point of view of coding theory.

Suppose that $N \subset \mathbb{F}_q^n$ is a Nikodym set. Recall that this means that for every point $x \in \mathbb{F}_q^n$, there is a line ℓ so that $\ell \setminus \{x\} \subset N$. The finite field Nikodym conjecture says that $|N| \geq c_n q^n$.

Nikodym sets have a close relationship with the Reed-Muller code. We consider a degree D in the range $nD < q - 1$. The Reed-Muller code takes as input an arbitrary function $g : \{0, ..., D\}^n \to \mathbb{F}_q$ and encodes it by extending it to a polynomial $P : \mathbb{F}_q^n \to \mathbb{F}_q$ with degree at most D in each coordinate. The key observation is that if we know P on a Nikodym set N, then we can recover the polynomial P everywhere, and hence we can recover the function g. For any point x, there is a line ℓ so that $\ell \setminus \{x\} \subset N$. We know the values of P on $\ell \setminus \{x\}$. The polynomial P has degree at most $nD < q - 1$, and so we can recover P on the line ℓ. In particular we can recover $P(x)$.

This shows that the Reed-Muller code gives an injective map from

$$\mathrm{Fcn}(\{0, ..., D\}^n, \mathbb{F}_q) \to \mathrm{Fcn}(N, \mathbb{F}_q).$$

This gives us a lower bound on the size of N:

$$|N| \geq |\{0, ..., D\}^n| = (D + 1)^n.$$

We are allowed to choose any D with $nD < q - 1$. This leads to the lower bound $|N| \geq c_n q^n$, and the constant c_n works out to be roughly n^{-n}. In this way, the Reed-Muller code proves the finite field Nikodym conjecture.

4.5. Conclusion and exercises

Polynomials over finite fields play an important role in error-correcting codes. Sudan's thesis [**Su**] refers to "the resilience of polynomials" (cf. the title of Chapter 2 of [**Su**]). The word resilience captures a property of polynomials that is important for coding theory: we can significantly distort the polynomial Q, but the information in Q survives. This resilience appears in several different ways, starting with the vanishing lemma, and it makes polynomials important tools in coding theory.

From the point of view of coding theory, we should mention one important caveat about the polynomial codes we have been discussing. The caveat is that the code is written as a string of symbols from $\mathbb{F}_q$, and it is a disadvantage for the code when q is large. The reason is that on a computer, information is stored as a series of 0's and 1's. An element of $\mathbb{F}_q$ could be stored as $\log_2 q$ bits. If even one of these bits changes, then the corresponding element in $\mathbb{F}_q$ changes also. Suppose we have recorded a list of elements of $\mathbb{F}_q$, and the data is corrupted. In order to guarantee that at most 49 % of the recorded elements of $\mathbb{F}_q$ are changed, we have to insist that at most a fraction $(49/100)(\log_2 q)^{-1}$ of the recorded bits are changed. If q is large, the amount of error we can tolerate at the level of bits gets smaller. The polynomial codes we have discussed require q to be large in order to be interesting. For the Reed-Solomon code, the message has length D and we require $q > 100D$. For the Reed-Muller code, the situation is a little better: the message has size D^n and we require $q > 100nD$. But even for Reed-Muller codes, q is still quite large.

On the other hand, polynomial codes have important applications in the theory of computational complexity, helping to understand the difficulty of finding approximate answers to computational problems. It is beyond the scope of this book to really explain these applications, but we can give the flavor of the subject by stating one of the results. As an example of a computational problem, we consider MAX-3SAT. We are given n Boolean variables $x_1, ...x_n$, and we are given a list of clauses of the form " NOT x_{i_1} and x_{i_2} and x_{i_3} ". (Each clause involves three of the variables, and each clause can have 0, 1, 2, or 3 NOTs arranged in any way. The length of the list of clauses is at most the number of possible clauses, which is $O(n^3)$.) Our job is to find the maximum number of clauses that can be satisfied by any choice of values for the Boolean variables $x_1, ..., x_n$. It is possible do this by brute force, checking all of the 2^n possible assignments to the variables $x_1, ..., x_n$. This process would take an exponential length of time. Is it possible to solve this problem in only polynomial time? This is (equivalent to) the famous $P \neq NP$ problem, and no one knows the answer. It sounds intuitive to almost all experts that there is no polynomial time algorithm to solve this problem.

Since no one knows how to solve MAX-3SAT efficiently, it is reasonable relax our goal and try to find an approximate solution instead of an exact solution. Instead of trying to find the exact maximum number of clauses that can be satisfied, what if we try to estimate this maximum number up to an error of 1%? If we allow a 1% error, is there a way to estimate the answer in much less time than it would take to try all the possibilities? This is a natural, fundamental question. Surprisingly, Arora, Lund, Motwani, Sudan, and Szegedy ([**ALMSS**]) proved that

this approximation problem is just as hard as the original problem! More precisely, if there were a polynomial-time algorithm to estimate MAX-3SAT up to 1% error, then it could be used as a subroutine to build a polynomial time algorithm to solve MAX-3SAT exactly. The proof in [**ALMSS**] uses polynomial codes in a crucial way. The proof method also applies to many other computational problems.

Polynomials (especially over finite fields) have a special structure which makes them useful tools in error-correcting codes and in complexity theory. Work in these fields has also led to new perspectives about polynomials.

For instance, here is an interesting problem about polynomials that plays a key role in the proof of the hardness of approximation theorem described in the last paragraph. This question could have been asked in the 1890s, but it wasn't asked until the 1990s, in connection with problems from computer science.

Suppose that $f : \mathbb{F}_q^n \to \mathbb{F}$ is a function. Suppose that on most lines in $\mathbb{F}_q^n$, the function f is close to a low-degree polynomial. To be more precise, suppose that for at least 99% of the lines $l \subset \mathbb{F}_q^n$, there is a polynomial P_l of degree at most d so that P_l agrees with f on at least 99% of the points of l. Does this imply that there is a polynomial $P : \mathbb{F}_q^n \to \mathbb{F}$ of degree at most d that agrees with f on at least 90% of the points of $\mathbb{F}_q^n$? The answer turns out to be yes as long as d is small compared to q (see Chapter 3 of [**Su**] and [**AS**]). If the dimension n is large, this theorem leads to a very efficient way to test whether a function f is close to a low degree polynomial: we take a few random lines l_j in $\mathbb{F}_q^n$, and we test whether the restriction of f to each line l_j is close to a polynomial. This test is surprisingly efficient because it is only necessary to look at f on a few lines, making up a tiny fraction of the points of $\mathbb{F}_q^n$.

We end the chapter with some exercises related to this theorem. The proof is based once again on the vanishing lemma and parameter counting, and the exercises give more practice using these tools. The main step in the proof of this n-dimensional theorem is a 2-dimensional result of Arora and Safra [**AS**].

THEOREM 4.9. (Arora-Safra) Suppose that $R, C : \mathbb{F}_q^2 \to \mathbb{F}_q$ are functions. For every $y \in \mathbb{F}_q$, $R_y(x) := R(x, y)$, and R_y is a polynomial in x of degree $\leq d$. Similarly, for every $x \in \mathbb{F}_q$, $C_x(y)$ is a polynomial in y of degree at most d. We also suppose that for a random point $(x, y) \in \mathbb{F}_q^2$,

$$\mathbb{P}[R(x, y) \neq C(x, y)] \leq \delta < \frac{1}{100}.$$

If $q > 10^6 d^2$, then there exists a polynomial $Q(x, y)$ with degree at most d in each variable so that

$$\mathbb{P}[Q(x, y) = R(x, y) = C(x, y)] \geq 1 - 10\delta.$$

We will break the proof of this theorem into a sequence of exercises. But before we start the details of the proof, let us discuss some of the issues and ideas. The first issue is, how can we find this polynomial $Q(x, y)$? Suppose that we find a subset $A \times B \subset \mathbb{F}_q^2$ with $|A| = |B| = d + 1$ so that $R(x, y) = C(x, y)$ for all $(x, y) \in A \times B$. Then there is a unique polynomial $Q(x, y)$ with degree at most d in each variable which agrees with R and C on $A \times B$. This Q is a reasonable guess for the polynomial we are looking for.

But will such a grid exist? Let $E \subset \mathbb{F}_q^2$ be the set of errors:

$$E := \{(x, y) \in \mathbb{F}_q^2 \text{ so that } R(x, y) \neq C(x, y)\}.$$

We wish to find a large grid $A \times B$ avoiding the set E. The size of E may be as large as $(1/100)q^2$ and d may be almost $q^{1/2}$. If we take a random subset of $(1/100)q^2$ points in $\mathbb{F}_q^2$, then the largest grid $A \times B$ avoiding this set of points will have $\min(|A|, |B|) \leq C \log q$, far smaller than d. So it doesn't sound likely that we will be able to find a $(d+1) \times (d+1)$ grid avoiding the set of errors.

However, the set of errors is not a random set. The hypotheses of the problem imply a lot of structure about the set of errors, and it turns out that there is indeed a $(10d) \times (10d)$ grid $A \times B$ without errors. Understanding the shape of the set of errors is a crucial step in the proof.

Given that the theorem is true, then it's not hard to see that there must be a very large grid without any errors. If we fix a value of y, then $Q(x, y)$ and $R(x, y)$ are each polynomials of degree at most d, and so either they agree for every value of x, or else they agree for at most d values of x. If the theorem is true, then $Q(x, y) = R(x, y)$ except for a bad set of at most $10\delta q^2$ points. Therefore, $Q(x, y) = R(x, y)$ except for a bad set of at most $20\delta q$ rows. Similarly, $Q(x, y) = C(x, y)$ except for a bad set of at most $20\delta q$ columns. Hence $R(x, y) = C(x, y)$ outside a small number of rows and a small number of columns.

But how can we use the hypotheses of the theorem in order to prove that there is a substantial error-free grid? Drawing on some ideas from the Berlekamp-Welch algorithm, Arora and Safra found a polynomial of the form

$$P(x, y, z) = P_0(x, y) + zP_1(x, y, z)$$

so that

- $\operatorname{Deg} P_0, \operatorname{Deg} P_1 \leq 10d$.
- There are subsets $G, H \subset \mathbb{F}_q$ with $|G|, |H| \geq (2/3)q$ so that for all $(x, y) \in G \times H$,

$$P(x, y, R(x, y)) = P(x, y, C(x, y)) = 0. \tag{4.1}$$

The proof that such a polynomial P exists uses the parameter counting argument and the vanishing lemma, roughly as in the Berlekamp-Welch algorithm. We use parameter counting to find a polynomial P so that Equation 4.1 holds at some points, and then we use the vanishing lemma to prove that it holds at many other points. Since $P(x, y, z) = P_0(x, y) + zP_1(x, y)$, Equation 4.1 implies that either $R(x, y) = C(x, y)$ or $P_1(x, y) = 0$. Therefore, we see that

$$R(x, y) = C(x, y) \text{ for all } (x, y) \in G \times H \setminus Z(P_1).$$

But $Z(P_1)$ is very small. It contains at most $10dq$ points. The number of errors in $\mathbb{F}_q^2$ may be as large as $(1/100)q^2$, but the number of errors in $G \times H$ is at most $10dq$. By hypothesis $q > 10^6 d^2$, and so $10dq$ is much smaller than $(1/100)q^2$. So we have found a large grid $G \times H$ where the density of errors is far smaller than in $\mathbb{F}_q^2$. There are so few errors in $G \times H$ that there has to be a $(10d) \times (10d)$ error-free sub grid $A \times B \subset G \times H$. Using this error-free grid, we can find the polynomial Q. Then more applications of the vanishing lemma show that $Q(x, y) = R(x, y) = C(x, y)$ for almost all $(x, y) \in \mathbb{F}_q^2$.

This finishes our outline of the proof of Theorem 4.9. Now we break the proof into a sequence of exercises. If you're interested in studying the proof, you can either go through the exercises, or try to prove Theorem 4.9 more independently, maybe looking at the exercises when you get stuck.

1. By an averaging argument, find sets $A_1, B_1 \subset \mathbb{F}_q$ with $|A_1| = |B_1| = 14d$ so that

$$\mathbb{P}_{(x,y)\in\mathbb{F}_q\times B_1}[R(x,y) \neq C(x,y)] \leq \delta.$$

2. By parameter counting, find a polynomial $P(x,y,z) = P_0(x,y) + zP_1(x,y)$ so that

- $P(x,y,R(x,y)) = 0$ for all $(x,y) \in A_1 \times B_1$.
- The polynomials P_0, P_1 have degree at most $10d$ in each variable x and y.

In the next steps, we use the vanishing lemma to show that $P(x,y,R(x,y))$ and $P(x,y,C(x,y))$ vanish at many other points $(x,y) \in \mathbb{F}_q^2$.

3. If $y \in B_1$ and $x \in \mathbb{F}_q$, show that

$$P(x,y,R(x,y)) = 0.$$

4. Let $G := \{x \in \mathbb{F} | R(x,y) = C(x,y)$ for at least $13d$ values of $y \in B_1\}$. By an averaging argument, show that

$$|G| \geq (2/3)q.$$

5. For every $x \in G, y \in \mathbb{F}_q$, show that

$$P(x,y,C(x,y)) = 0.$$

6. Let $H = \{y \in \mathbb{F}_q | R(x,y) = C(x,y)$ for at least $13d$ values of $x \in G\}$. By an averaging argument, show that

$$|H| \geq (2/3)q.$$

7. For all $y \in H, x \in \mathbb{F}_q$, show that

$$P(x,y,R(x,y)) = 0.$$

Summarizing what we did so far, we have found a large grid $G \times H$ so that for all $(x,y) \in G \times H$,

$$P(x,y,R(x,y)) = P(x,y,C(x,y)) = 0.$$

Since $P(x,y,z) = P_0(x,y) + zP_1(x,y)$, Equation 4.1 implies that either $R(x,y) = C(x,y)$ or $P_1(x,y) = 0$. Therefore, we see that

$$R(x,y) = C(x,y) \text{ for all } (x,y) \in G \times H \setminus Z(P_1).$$

But $Z(P_1)$ is very small. It contains at most $10dq$ points.

8. Prove that there are subsets $A \subset G$ and $B \subset H$ with $|A| = |B| = 10d$ so that $A \times B$ contains no points of E and so that

$$\mathbb{P}_{(x,y)\in\mathbb{F}_q\times B}[R(x,y) \neq C(x,y)] \leq 2\delta.$$

9. Show that there is a polynomial $Q(x,y)$ with degree at most d in each variable so that $Q(x,y) = R(x,y) = C(x,y)$ on $A \times B$.

Now we again use the vanishing lemma to show that $Q(x,y) = R(x,y)$ and/or $Q(x,y) = C(x,y)$ at many other points $(x,y) \in \mathbb{F}_q^2$.

10. Show that $Q(x,y) = R(x,y)$ on $\mathbb{F}_q \times B$.

11. Show that there is a subset $G' \subset \mathbb{F}_q$ with $|G'| \geq (1-3\delta)q$ so that $Q(x,y) = C(x,y)$ on $G' \times \mathbb{F}_q$.

12. Show that there is a subset $H' \subset \mathbb{F}_q$ with $|H'| \geq (1-3\delta)q$ so that $Q(x,y) = R(x,y)$ on $\mathbb{F}_q \times H'$.

Finally, we see that $Q(x, y) = R(x, y) = C(x, y)$ on $G' \times H'$, which has size at least $(1 - 10\delta)q^2$.

More results in this direction and the connection to hardness of approximation are explained in [**Su**], [**AS**], and [**ALMSS**].

https://doi.org/10.1090//ulect/064/05

CHAPTER 5

On polynomials and linear algebra in combinatorics

There is a lot of interesting work in the combinatorics literature that involves polynomials and linear algebra in an indirect way. We mentioned some directions in the introduction, in Section 1.4. In this chapter, we describe in detail one proof that has a similar flavor to the proof of finite field Kakeya. At the end of the chapter, we will give some references to further reading about these ideas.

The theorem we prove in this chapter is about the distinct distance problem in high dimensions. One of the main problems in the book is the distinct distance problem in the plane: given a set of N points in the plane, what is the smallest possible number of distinct distances determined by the set. It is also natural to ask about the same problem in higher dimensions. In the 1970s, Larman, Rogers, and Seidel [**LRS77**] proved the following theorem about the regime where the number of distinct distances is fixed and the number of dimensions goes to infinity.

Recall that we write $d(P)$ for the set of (non-zero) distances among points of P:

$$d(P) := \{|p_1 - p_2|\}_{p_1,p_2 \in P; p_1 \neq p_2}.$$

THEOREM 5.1. ([**LRS77**]) Suppose that $P \subset \mathbb{R}^n$ is a set, and that $|d(P)| \leq s$. Then

$$|P| \leq \binom{n+s+1}{s}.$$

On the other hand, there is an example of a set $P \subset \mathbb{R}^n$ with $|d(P)| = s$ and

$$|P| = \binom{n+1}{s}.$$

If we fix s and send $n \to \infty$, the ratio between the upper bound and the lower bound approaches 1. In this regime, the theorem is very accurate. The case $s = 2$ is already interesting. On the other hand, if we take $n = 2$, then we get the bound $|P| \leq \binom{s+3}{3} \sim s^3$. In other words, this gives the lower bound $|d(P)| \gtrsim |P|^{1/3}$. More generally, if n is fixed and $s \to \infty$, the bounds here are not very good.

PROOF. We start with the example. For any subset $A \subset \{1, ..., n+1\}$, we define the point $p(A) \in \mathbb{R}^{n+1}$ to be the vector $(p_1, ..., p_{n+1})$ with $p_j = 1$ if $j \in A$ and $p_j = 0$ if $j \notin A$. We let P be the set of points $p(A)$ given by all s-element subsets $A \subset \{1, ..., n+1\}$. We have $|P| = \binom{n+1}{s}$. The distance between $p(A_1)$ and $p(A_2)$ only depends on the cardinality of $A_1 \cap A_2$. If $A_1 \neq A_2$, then $0 \leq |A_1 \cap A_2| \leq s-1$, and so $|d(P)| = s$. The set P has been written as a subset of $\mathbb{R}^{n+1}$, but it actually lies in the n-dimensional hyperplane defined by $x_1 + ... + x_{n+1} = s$. Therefore, we can consider P as a subset of $\mathbb{R}^n$. This finishes the discussion of the example.

Now we turn to the upper bound. Suppose that P is the set $\{p_1, ..., p_N\}$, with $N = |P|$. For every point $p_j \in P$, we construct a polynomial f_j so that $f_j(p_j) \neq 0$ but $f_j(p_{j'}) = 0$ for all $j' \neq j$. To construct f_j, suppose that $d_1, ..., d_s$ are the distances in $d(P)$. Recall that $d(P)$ is the set of non-zero distances between points of P, so none of the d_r is zero. Now we define the polynomial f_j by

$$f_j(x) = \prod_{r=1}^{s} \left(|x - p_j|^2 - d_r^2 \right).$$

For any $j' \neq j$, we have $|p'_j - p_j| = d_r$ for some r, and so $f_j(p_{j'}) = 0$. On the other hand, $f_j(p_j) = \prod_{r=1}^{s}(-d_r^2) \neq 0$.

Because $f_j(p_{j'}) \neq 0$ if $j = j'$ and 0 if $j \neq j'$, it follows that the functions f_j are linearly independent. Indeed, suppose that there are some real numbers λ_j so that the function $\sum_{j=1}^{N} \lambda_j f_j$ is equal to zero. Evaluating this function at the point p_j, we see that $\lambda_j = 0$. This holds for every j, so all the λ_j vanish. This shows that the functions f_j are indeed linearly independent.

Now the f_j are all polynomials of degree at most $2s$. Since they are linearly independent in $\mathrm{Poly}_{2s}(\mathbb{R}^n)$, we get the bound

$$N \leq \mathrm{Dim}\, \mathrm{Poly}_{2s}(\mathbb{R}^n) = \binom{n+2s}{n}.$$

But we can get a better bound than this, because the span of the polynomials f_j lies in a subspace of $\mathrm{Poly}_{2s}(\mathbb{R}^n)$. To find this subspace, we expand everything out in coordinates: we let $x = (x_1, ..., x_n)$ and we let $p_j = (p_{j,1}, ..., p_{j,n})$. We use the index k to label the coordinates. First we expand out $|x - p_j|^2$.

$$|x - p_j|^2 = \sum_{k=1}^{n} (x_k - p_{j,k})^2 = \sum_{k=1}^{n} \left(x_k^2 - 2p_{j,k}x_k + p_{j,k}^2 \right).$$

So we see that $|x - p_j|^2 - d_r^2$ is in the span of the following polynomials:

$$1, x_1, ..., x_n, \left(\sum_{k=1}^{n} x_k^2 \right).$$

Define $g_0 = 1$, $g_j = x_j$ for $j = 1, ...n$, and $g_{n+1} = \sum_k x_k^2$. Every function f_j can be written as a homogeneous polynomial in the g_j (where $j = 0, ..., n+1$) of degree s. The dimension of the space of homogeneous polynomials in the g_j of degree s is $\binom{n+1+s}{s}$. Therefore, the span of the functions f_j lies in a subspace of $\mathrm{Poly}_{2s}(\mathbb{R}^n)$ of dimension at most $\binom{n+s+1}{s}$. □

The key idea in the proof, using linear independence to prove that the number of functions f_j is not too big, goes back to a slightly earlier paper by Koornwinder [**Koo**]. This idea has many more applications. There are two very engaging books that explain the further developments of this idea. One is the book *Linear algebra methods in combinatorics*, by Babai and Frankl, [**BF**]. This book was never quite finished, but there is a draft available on Babai's webpage. In spite of being not quite finished, it is very clearly written. A second book that discusses many of the same ideas is *Thirty three miniatures* by Matousek, [**Ma2**]. There is also a literature on blocking numbers which involves related arguments – for instance, see the short paper [**BrSc**] by Brouwer and Schrijver.

There are some similarities and also some differences between this argument and the proof of finite field Kakeya. Both arguments use linear algebra on the space of polynomials, and the flavor of the questions is similar. On the other hand, the rank-nullity theorem is used in the opposite direction in some sense. In finite field Kakeya, we use dimension arguments to prove that there must exist a polynomial with certain properties, even though it's hard to write down the polynomial. In this proof, we write down some polynomials f_j fairly explicitly and we use dimension arguments to show that there are not too many polynomials in our list.

https://doi.org/10.1090//ulect/064/06

CHAPTER 6

The Bezout theorem

As we continue to develop the polynomial method, we will need to use the Bezout theorem, a fundamental result from algebraic geometry. We will give a proof that connects to some of the ideas we have seen in previous chapters. In particular, the dimension of the space of polynomials $\mathrm{Poly}_D(\mathbb{F}^n)$ will be an important character in the argument.

The Bezout theorem controls the intersection of two algebraic varieties. There are many variations of the Bezout theorem, and we will only discuss a couple. Here is the simplest version of the theorem.

THEOREM 6.1. (Bezout in the plane) Suppose $\mathbb{F}$ is a field and P, Q are polynomials in $\mathrm{Poly}(\mathbb{F}^2)$ with no common factor. Let $Z(P,Q) := \{(x,y) \in \mathbb{F}^2 | P(x,y) = Q(x,y) = 0\}$. Then the number of points in $Z(P,Q)$ is at most $(\mathrm{Deg}\, P)(\mathrm{Deg}\, Q)$.

We recall some basic facts about factorization in a polynomial ring $\mathrm{Poly}(\mathbb{F}^n)$. The units in $\mathrm{Poly}(\mathbb{F}^n)$ are the non-zero elements of $\mathbb{F}$. We say that P and Q have no common factor if every common factor of P and Q is a unit. We say that P is reducible if $P = Q_1 \cdot Q_2$ where neither of Q_1, Q_2 is a unit. Otherwise, we say that P is irreducible.

Polynomial rings have unique factorization (cf. Corollary 2.4 in Chapter 4 of [**Lan**]). In particular, the following proposition holds.

PROPOSITION 6.2. If $P, Q, R \in \mathrm{Poly}(\mathbb{F}^n)$ and P divides $Q \cdot R$, and if P and Q have no common factor, then P divides R.

There are several approaches to proving the Bezout theorem. I found the approach that we use here in Joe Harris's book *Algebraic Geometry, a First Course*, [**Ha**], Exercise 13.17.

6.1. Proof of the Bezout theorem

We recall that $\mathrm{Poly}_D(\mathbb{F}^n)$ is the vector space of polynomials of degree at most D. Suppose that $I \subset \mathrm{Poly}(\mathbb{F}^n)$ is an ideal. We define I_D to be $I \cap \mathrm{Poly}_D(\mathbb{F}^n)$. We will also be interested in the ring $R := \mathrm{Poly}(\mathbb{F}^n)/I$. We define the vector space $R_D \subset R$ to be $\mathrm{Poly}_D(\mathbb{F}^n)/I_D$. In other words, an element of R lies in R_D if and only if it can be represented by a polynomial $P \in \mathrm{Poly}_D(\mathbb{F}^n)$. The spaces R_D are nested: $R_0 \subset R_1 \subset ... \subset R$. The dimensions of I_D and R_D for various ideals I will be the main characters in the proof of the Bezout theorem.

We begin with some lemmas that explain how the number of points in a set like $Z(P,Q)$ is connected to the dimensions of ideals.

LEMMA 6.3. If $X \subset \mathbb{F}^n$ is any finite set, and $f : X \to \mathbb{F}$ is any function, then there is a polynomial P of degree $\leq |X| - 1$ which agrees with f on X.

PROOF. For each $p \in X$, we will construct a polynomial P_p with $P_p(p) = 1$ and $P_p = 0$ on $X \setminus p$. Fix p. For each $q \in X \setminus p$, let L_q be a polynomial that vanishes at q but not at p. Then define $P_p = c \prod_{q \in X \setminus p} L_q$. We see that $P_p(q) = 0$ for each $q \in X \setminus p$, and that $P_p(p) \neq 0$. By choosing the constant c, we can arrange that $P_p(p) = 1$. The degree of P_p is $|X| - 1$.

Finally, for an arbitrary function f, we define $P = \sum_{p \in X} f(p) P_p$. □

LEMMA 6.4. Let $I \subset \mathrm{Poly}(\mathbb{F}^n)$ be an ideal, and let $R := \mathrm{Poly}(\mathbb{F}^n)/I$. We write $\mathrm{Dim}\, R$ for the dimension of R as a vector space over $\mathbb{F}$, which may be finite or infinite. Then $|Z(I)| \leq \mathrm{Dim}\, R$.

PROOF. For any set $X \subset \mathbb{F}^n$, let $E_X : \mathrm{Poly}(\mathbb{F}^n) \to \mathrm{Fcn}(X, \mathbb{F})$ be the map that restricts each polynomial $P \in \mathrm{Poly}(\mathbb{F}^n)$ to a function on X. If $X \subset Z(I)$, then $I \subset \mathrm{Ker}\, E_X$, and so E_X descends to a map $E_X : R \to \mathrm{Fcn}(X, \mathbb{F})$.

If $X \subset Z(I)$ is any finite set, then Lemma 6.3 says that E_X is surjective. Therefore, $\mathrm{Dim}\, R \geq |X|$ for any finite subset $X \subset Z(I)$. □

Now we can begin the proof of the Bezout theorem, Theorem 6.1

PROOF. Suppose that $P, Q \in \mathrm{Poly}(\mathbb{F}^2)$ have no common factor. To prove the Bezout theorem, we will apply Lemma 6.4 to the ideal $I = (P, Q)$, giving the inequality:

$$|Z(P, Q)| \leq \mathrm{Dim}\, R = \mathrm{Dim}\left(\mathrm{Poly}(\mathbb{F}^2)/(P, Q)\right).$$

So to prove the Bezout theorem, it suffices to show that for arbitrarily large degrees D,

$$\mathrm{Dim}\left(\mathrm{Poly}_D(\mathbb{F}^2)/(P, Q)_D\right) \leq (\mathrm{Deg}\, P)(\mathrm{Deg}\, Q). \tag{6.1}$$

To estimate this dimension, we consider the following sequence of quotient maps:

$$\mathrm{Poly}_D(\mathbb{F}^2) \xrightarrow{\alpha} \mathrm{Poly}_D(\mathbb{F}^2)/(P)_D \xrightarrow{\beta} \mathrm{Poly}_D(\mathbb{F}^2)/(P, Q)_D.$$

We want to estimate the dimension of the last space on the right. We start with the dimension of the first space on the left and estimate the dimensions of all the kernels and images.

The dimension of $\mathrm{Poly}_D(\mathbb{F}^n)$ has played a fundamental role in earlier chapters. Near the beginning of the book, in Lemma 2.2, we calculated $\mathrm{Dim}\, \mathrm{Poly}_D(\mathbb{F}^n)$ for any n. In particular, we saw that

$$\mathrm{Dim}\, \mathrm{Poly}_D(\mathbb{F}^2) = \binom{D+2}{2}. \tag{6.2}$$

The kernel of the map α is $(P)_D$. Next we calculate its dimension.

LEMMA 6.5. If $P \in \mathrm{Poly}(\mathbb{F}^2)$ is a non-zero polynomial, then for all $D \geq \mathrm{Deg}\, P$,

$$\mathrm{Dim}(P)_D = \mathrm{Dim}\, \mathrm{Poly}_{D - \mathrm{Deg}\, P}(\mathbb{F}^2) = \binom{D - \mathrm{Deg}\, P + 2}{2}.$$

PROOF. Define μ_P by $\mu_P(R) = PR$. The map μ_P is a linear map from $\mathrm{Poly}_{D - \mathrm{Deg}\, P}$ to $(P)_D$. We claim this linear map is an isomorphism. The kernel of the map is zero. Any element in $(P)_D$ can be written as PR for some $R \in \mathrm{Poly}(\mathbb{F}^2)$. We have $D \geq \mathrm{Deg}(PR) = \mathrm{Deg}\, P + \mathrm{Deg}\, R$, and so $\mathrm{Deg}\, R \leq D - \mathrm{Deg}\, P$ and $R \in \mathrm{Poly}_{D - \mathrm{Deg}\, P}(\mathbb{F}^2)$. This shows that μ_P is surjective. Therefore, μ_P is an isomorphism and the dimension of its domain equals the dimension of its range. □

The map α is clearly surjective, and so we get

$$\operatorname{Dim}\left(\operatorname{Poly}_D(\mathbb{F}^2)/(P)_D\right) = \operatorname{Dim}\operatorname{Image}\alpha = \operatorname{Dim}\operatorname{Poly}_D(\mathbb{F}^2) - \operatorname{Dim}\operatorname{Ker}\alpha.$$

Plugging in our results for these dimensions, we get

$$\begin{aligned}\operatorname{Dim}\left(\operatorname{Poly}_D(\mathbb{F}^2)/(P)_D\right) &= \operatorname{Dim}\operatorname{Poly}_D(\mathbb{F}^2) - \operatorname{Dim}(P)_D \\ &= \binom{D+2}{2} - \binom{D-\operatorname{Deg}P+2}{2}.\end{aligned}$$

For all $D \geq \operatorname{Deg}P$, we can expand $\binom{D-\operatorname{Deg}P+2}{2}$ and $\binom{D}{2}$ to get

$$\operatorname{Dim}\left(\operatorname{Poly}_D(\mathbb{F}^2)/(P)_D\right) = (\operatorname{Deg}P)D + (1/2)(3\operatorname{Deg}P - (\operatorname{Deg}P)^2)$$

The right-hand side is a degree 1 polynomial in D. The constant term is a little complicated. The exact form of the constant term is not important for our argument, so we abbreviate it as $c(P)$:

$$\operatorname{Dim}\left(\operatorname{Poly}_D(\mathbb{F}^2)/(P)_D\right) = (\operatorname{Deg}P)D + c(P). \tag{6.3}$$

Now we study the map β, which is a little bit subtler than α. We estimate the dimension of $\operatorname{Ker}\beta$.

LEMMA 6.6.

$$\operatorname{Dim}\operatorname{Ker}\beta \geq \operatorname{Dim}\left(\operatorname{Poly}_{D-DegQ}(\mathbb{F}^2)/(P)_{D-DegQ}\right).$$

PROOF. Let μ_Q be the linear map $\mu_Q(R) := QR$. The map μ_Q is well-defined on various spaces. In particular,

$$\mu_Q : \operatorname{Poly}_{D-\operatorname{Deg}Q}(\mathbb{F}^2) \to \operatorname{Poly}_D(\mathbb{F}^2).$$

$$\mu_Q : (P)_{D-\operatorname{Deg}Q} \to (P)_D.$$

Therefore, μ_Q descends to the quotient:

$$\mu_Q : \operatorname{Poly}_{D-\operatorname{Deg}Q}(\mathbb{F}^2)/(P)_{D-\operatorname{Deg}Q} \to \operatorname{Poly}_D(\mathbb{F}^2)/(P)_D.$$

From now on, we fix this domain and range for μ_Q. The image of μ_Q lies in the kernel of β. Using the fact that P and Q have no common factor, we will prove that μ_Q is injective. This will imply that

$$\operatorname{Dim}\operatorname{Ker}\beta \geq \operatorname{Dim}\operatorname{Image}\mu_Q = \operatorname{Dim}\left(\operatorname{Poly}_{D-\operatorname{Deg}Q}(\mathbb{F}^2)/(P)_{D-\operatorname{Deg}Q}\right).$$

It just remains to prove μ_Q is injective. Suppose $r \in \operatorname{Ker}\mu_Q \subset \operatorname{Poly}_{D-\operatorname{Deg}Q}(\mathbb{F}^2)/(P)_{D-\operatorname{Deg}Q}$. Let $R \in \operatorname{Poly}_{D-\operatorname{Deg}Q}(\mathbb{F}^2)$ be a polynomial representing r. We know that QR vanishes in $\operatorname{Poly}_D(\mathbb{F}^2)/(P)_D$, and so QR lies in (P). Since P and Q have no common factor, Proposition 6.2 implies that $R \in (P)$. But then $r = 0$ in $\operatorname{Poly}_{D-\operatorname{Deg}Q}(\mathbb{F}^2)/(P)_{D-\operatorname{Deg}Q}$. This shows that μ_Q is injective, finishing the proof. □

Remark: It is not always true that

$$\operatorname{Dim}\operatorname{Ker}\beta = \operatorname{Dim}\left(\operatorname{Poly}_{D-\operatorname{Deg}Q}(\mathbb{F}^2)/(P)_{D-\operatorname{Deg}Q}\right).$$

It is a good exercise for the reader to find a counterexample.

The map β is surjective, and so we see that

$$\operatorname{Dim}\left(\operatorname{Poly}_D(\mathbb{F}^2)/(P,Q)_D\right) = \operatorname{Dim}\operatorname{Image}\beta = \operatorname{Dim}\left(\operatorname{Poly}_D(\mathbb{F}^2)/(P)_D\right) - \operatorname{Dim}\operatorname{Ker}\beta.$$

Plugging in Lemma 6.6, we get

$$\begin{aligned}\operatorname{Dim}\left(\operatorname{Poly}_D(\mathbb{F}^2)/(P,Q)_D\right) \leq \operatorname{Dim}&\left(\operatorname{Poly}_D(\mathbb{F}^2)/(P)_D\right)\\ &- \operatorname{Dim}\left(\operatorname{Poly}_{D-\operatorname{Deg} Q}(\mathbb{F}^2)/(P)_{D-\operatorname{Deg} Q}\right).\end{aligned}$$

Now if D is sufficiently large, we can plug in Equation 6.3 to the right-hand side, getting

$$\begin{aligned}\operatorname{Dim}\left(\operatorname{Poly}_D(\mathbb{F}^2)/(P,Q)_D\right) &\leq [(\operatorname{Deg} P)D + c(P)] - [(\operatorname{Deg} P)(D - \operatorname{Deg} Q) + c(P)]\\ &= (\operatorname{Deg} P)(\operatorname{Deg} Q).\end{aligned}$$

This proves Inequality 6.1 and finishes the proof of the Bezout theorem. □

6.2. A Bezout theorem about surfaces and lines

To study lines in 3-dimensional space, we will need the following variation of the Bezout theorem.

THEOREM 6.7. Suppose that $\mathbb{F}$ is an infinite field. If $P, Q \in \operatorname{Poly}(\mathbb{F}^3)$ have no common factor, then the number of lines in $Z(P,Q) \subset \mathbb{F}^3$ is at most $(\operatorname{Deg} P)(\operatorname{Deg} Q)$.

We approach this result by generalizing the proof of Theorem 6.1.

PROOF. We define I to be the ideal generated by P and Q, and we define R to be the ring $\operatorname{Poly}(\mathbb{F}^3)/I$. As above, define $I_D := \operatorname{Poly}_D(\mathbb{F}^3) \cap I$ to be the polynomials in I of degree at most D. Define $R_D \subset R$ to be the elements of R that can be represented by a polynomial of degree at most D. We have $R_D = \operatorname{Poly}_D(\mathbb{F}^3)/I_D$.

On the one hand, we will bound the dimension of R_D from above in terms of the degrees of P and Q:

$$\operatorname{Dim} R_D \leq (\operatorname{Deg} P)(\operatorname{Deg} Q)D + c(P,Q). \tag{6.4}$$

On the other hand, if $Z(P,Q)$ contains L lines, then we will bound the dimension of R_D from below as follows:

$$\operatorname{Dim} R_D \geq LD - c(L). \tag{6.5}$$

Given these two bounds, taking $D \to \infty$, we see that $L \leq (\operatorname{Deg} P)(\operatorname{Deg} Q)$. Therefore, it suffices to establish these two inequalities.

We begin with the upper bound on $\operatorname{Dim} R_D$, Equation 6.4. We closely follow the argument in the planar case. We have $R_D = \operatorname{Poly}_D(\mathbb{F}^3)/(P,Q)_D$. We study the sequence of quotient maps

$$\operatorname{Poly}_D(\mathbb{F}^3) \xrightarrow{\alpha} \operatorname{Poly}_D(\mathbb{F}^3)/(P)_D \xrightarrow{\beta} \operatorname{Poly}_D(\mathbb{F}^3)/(P,Q)_D.$$

By Lemma 2.2,

$$\operatorname{Dim} \operatorname{Poly}_D(\mathbb{F}^3) = \binom{D+3}{3}. \tag{6.6}$$

By the same argument as Lemma 6.5, the dimension of $(P)_D$ is equal to $\operatorname{Dim} \operatorname{Poly}(\mathbb{F}^3)_{D-\operatorname{Deg} P}$. For $D \geq \operatorname{Deg} P$, this is equal to $\binom{D-\operatorname{Deg} P+3}{3}$. In this range of D, we get:

$$\begin{aligned}\operatorname{Dim}\left(\operatorname{Poly}_D(\mathbb{F}^3)/(P)_D\right) &= \operatorname{Dim} \operatorname{Poly}_D(\mathbb{F}^3) - \operatorname{Dim}(P)_D\\ &= \binom{D+3}{3} - \binom{D-\operatorname{Deg} P+3}{3}.\end{aligned}$$

The right-hand side is a polynomial in D of degree 2. The coefficient of D^2 is $(1/2)(\operatorname{Deg} P)$. The lower coefficients depend on $\operatorname{Deg} P$ in a more complicated way, but they don't play a role in our proof. Focusing on the highest-order term, we write

$$\operatorname{Dim}\left(\operatorname{Poly}_D(\mathbb{F}^3)/(P)_D\right) = (1/2)(\operatorname{Deg} P)D^2 + c_1(P)D + c_0(P). \tag{6.7}$$

By the same argument as Lemma 6.6, we see that

$$\operatorname{Dim}\operatorname{Ker}\beta \geq \operatorname{Dim}\left(\operatorname{Poly}_{D-\operatorname{Deg} Q}(\mathbb{F}^3)/(P)_{D-\operatorname{Deg} Q}\right).$$

Since β is surjective, we see that

$$\operatorname{Dim}\left(\operatorname{Poly}_D(\mathbb{F}^3)/(P,Q)_D\right) = \operatorname{Dim}\operatorname{Image}\beta = \operatorname{Dim}\left(\operatorname{Poly}_D(\mathbb{F}^3)/(P)_D\right) - \operatorname{Dim}\operatorname{Ker}\beta \leq$$

$$\leq \operatorname{Dim}\left(\operatorname{Poly}_D(\mathbb{F}^3)/(P)_D\right) - \operatorname{Dim}\left(\operatorname{Poly}_{D-\operatorname{Deg} Q}(\mathbb{F}^3)/(P)_{D-\operatorname{Deg} Q}\right).$$

Plugging in Equation 6.7, we get

$$\operatorname{Dim}\left(\operatorname{Poly}_D(\mathbb{F}^3)/(P,Q)_D\right) \leq (\operatorname{Deg} P)(\operatorname{Deg} Q)D + c(P,Q).$$

This finishes the proof of Equation 6.4.

Now we turn to the lower bounds on the size of R_D related to the lines in $Z(P,Q)$.

For any set $X \subset \mathbb{F}^n$, let E_X be the restriction map from $\operatorname{Poly}_D(\mathbb{F}^n)$ to $\operatorname{Fcn}(X,\mathbb{F})$.

LEMMA 6.8. If $\mathbb{F}$ is an infinite field, and if X is a union of L lines in $\mathbb{F}^n$, then the rank of $E_X : \operatorname{Poly}_D \to \operatorname{Fcn}(X,\mathbb{F})$ is $\geq LD - c(L)$.

PROOF. Fix D. After a linear change of variables, we can assume that each line is transverse to planes of the form $x_n = h$. Choose $D - L$ values $h_1, ..., h_{D-L}$ so that each plane $x_n = h_j$ intersects the L lines in L distinct points. Let $X_0 \subset X$ be these $L(D-L)$ points.

We claim that for any function $f : X_0 \to \mathbb{F}$, there is a degree D polynomial that agrees with f on X_0. This will imply that $\operatorname{Rank} E_X : \operatorname{Poly}_D(\mathbb{F}^n) \to \operatorname{Fcn}(X,\mathbb{F})$ is at least $|X_0| = LD - L^2$.

Fix a value h_j. The set X_0 intersects the plane $x_n = h_j$ at L points. Call these points $(y_{1,j}, h_j), ..., (y_{L,j}, h_j)$ where $y_{k,j} \in \mathbb{F}^{n-1}$. By Lemma 6.3, we can find a polynomial $P_j \in \operatorname{Poly}_L(\mathbb{F}^{n-1})$ so that $P_j(y_{k,j}) = f(y_{k,j})$ for each $1 \leq k \leq L$.

Now we want to find a polynomial $P \in \operatorname{Poly}_D(\mathbb{F}^n)$ so that $P(y, h_j) = P_j(y)$ for all y and all j from 1 to $D - L$. Let's expand out P_j and P:

$$P_j(y) = \sum_I c_I(j) y^I,$$

where I is a multi-index in $(n-1)$ variables of degree at most L.

Now we will choose P to have the following form:

$$P(y, x_n) = \sum_I P_I(x_n) y^I, \text{ where } |I| \leq L \text{ and } \operatorname{Deg} P_I \leq D - L.$$

It suffices to choose P_I so that $P_I(h_j) = c_I(j)$ for each $j = 1, ..., D - L$. We can do this by applying Lemma 6.3 again. □

Suppose that $Z(P,Q)$ contains L lines, and let X be the union of these lines. The dimension of R_D is at least the rank of $E_X : R_D \to \operatorname{Fcn}(X,\mathbb{F})$. By the last

Lemma, this rank is at least $DL - L^2$. Therefore, $\operatorname{Dim} R_D \geq DL - L^2$, proving Inequality 6.5.

This finishes the proof of Theorem 6.7. □

EXERCISE 6.1. To what extent does Theorem 6.7 hold in finite fields?

6.3. Hilbert polynomials

In the proof of the Bezout theorem, we studied the dimension of vector spaces of the form

$$(\operatorname{Poly}(\mathbb{F}^n)/I)_D := \operatorname{Poly}_D(\mathbb{F}^n)/I_D.$$

These dimensions have been studied a lot in algebraic geometry, and they also play an important role in the polynomial arguments in this book. The goal of this section is to give a very brief introduction to the study of these dimensions in algebraic geometry, with references to some further reading.

One fundamental theorem says that for large values of D, the dimension of the space $\operatorname{Poly}_D(\mathbb{F}^n)/I_D$ is given by a polynomial in D, called an affine Hilbert polynomial. (A small variation of this result was proven by Hilbert.)

PROPOSITION 6.9. If $I \subset \operatorname{Poly}(\mathbb{F}^n)$ is an ideal, then there exists a polynomial $h_{\operatorname{Poly}(\mathbb{F}^n)/I}(D)$ and a number D_0, so that for all $D \geq D_0$,

$$\operatorname{Dim}(\operatorname{Poly}(\mathbb{F}^n)/I)_D = h_{\operatorname{Poly}(\mathbb{F}^n)/I}(D).$$

Chapter 1.9 of [**Ei**] gives a nice introduction to Hilbert polynomials. I couldn't find a nice reference for the exact statement of Proposition 6.9, but it does follow quickly from a general theorem in [**Ei**]. At the end of the section, we quote Theorem 1.11 from [**Ei**] and explain how to deduce Proposition 6.9 from it.

We give a couple examples of Hilbert polynomials from the arguments we have seen. In the first chapter, we computed the dimension of $\operatorname{Poly}_D(\mathbb{F}^n)$. If the ideal I is zero, then for all $D \geq 0$, we see that

$$\operatorname{Dim} \operatorname{Poly}_D(\mathbb{F}^n)/I_D = \operatorname{Dim} \operatorname{Poly}_D(\mathbb{F}^n) = \binom{D+n}{n} = (1/n!)D^n + ... \tag{6.8}$$

In the proof of the Bezout theorem in the plane, we considered the principal ideal $(P) \subset \operatorname{Poly}(\mathbb{F}^2)$. We proved that for all $D \geq \operatorname{Deg} P$,

$$\operatorname{Dim}(\operatorname{Poly}(\mathbb{F}^2)/(P))_D = (\operatorname{Deg} P)D + (1/2)(3 \operatorname{Deg} P - (\operatorname{Deg} P)^2). \tag{6.9}$$

If $I \subset \operatorname{Poly}(\mathbb{F}^n)$ is an ideal, then the Hilbert polynomial $h_{\operatorname{Poly}(\mathbb{F}^n)/I}(D)$ encodes some information about the variety $Z(I)$ defined by the ideal I. In particular, the dimension of $Z(I)$ can be defined to be the degree of the Hilbert polynomial $h_{\operatorname{Poly}(\mathbb{F}^n)/I}(D)$. The reader can check this in the two examples above. When I is the zero ideal in $\operatorname{Poly}(\mathbb{F}^n)$, then $Z(I) = \mathbb{F}^n$, which should have dimension n. The degree of the Hilbert polynomial is indeed n. When I is the ideal $(P) \subset \operatorname{Poly}(\mathbb{F}^2)$, then $Z(I)$ is a curve $\mathbb{F}^2$, and it should have dimension 1. The degree of the Hilbert polynomial is indeed 1.

This discussion ties back to the polynomial method. In polynomial method proofs, like the proof of finite field Kakeya, it is crucial that the dimension of $\operatorname{Poly}_D(\mathbb{F}^n)$ grows like D^n. In other words, it is crucial that $h_{\operatorname{Poly}(\mathbb{F}^n)}(D)$ has degree n. According to the point of view of Hilbert polynomials, this is a way of saying that $\mathbb{F}^n$ has dimension n.

For more information about the dimensions of varieties and about Hilbert polynomials, the reader can consult [**Ha**] and [**Ei**].

To finish the Section, for completeness, we explain how to reduce Proposition 6.9 to Theorem 1.11 in [**Ei**]. Theorem 1.11 is a general theorem saying that certain dimension functions are polynomials. It is stated in terms of graded modules. Let M be a finitely generated module over the ring $\mathrm{Poly}(\mathbb{F}^n) = \mathbb{F}[x_1, ..., x_n]$. We say that M is a graded module if $M = \bigoplus_{D \geq 0} M_D$ and if, for any homogeneous polynomial P of degree E, and any $m \in M_E$, $Pm \in M_{D+E}$. Theorem 1.11 in [**Ei**] says that if M is a finitely generated graded module over $\mathrm{Poly}(\mathbb{F}^n)$, then there is a polynomial $h_M(D)$ and an integer D_0 so that $\mathrm{Dim}\, M_D = h_M(D)$ for all $D \geq D_0$. We reduce Proposition 6.9 to this result as follows. Let $I \subset \mathrm{Poly}(\mathbb{F}^n)$ be an ideal. Let $I_D \subset I$ be $I \cap \mathrm{Poly}_D(\mathbb{F}^n)$. Let $R = \mathrm{Poly}(\mathbb{F}^n)/I$, and let $R_D = \mathrm{Poly}_D(\mathbb{F}^n)/I_D$. We have $R_0 \subset R_1 \subset R_2 \subset ....$ Next we define $M_D := R_D/R_{D-1}$. We let $M = \bigoplus_{D \geq 0} M_D$. We claim that M is a finitely generated graded module over $\mathrm{Poly}(\mathbb{F}^n)$. If P is a homogeneous polynomial of degree E, and $m \in M_D$ is represented by a polynomial $Q \in \mathrm{Poly}_D(\mathbb{F}^n)$, then we define Pm to be the class of PQ in $M_{D+E} = R_{D+E}/R_{D+E-1}$. We can assume that I is a proper ideal, so $I_0 = 0$, and so $M_0 = R_0 = \mathbb{F}$. The element $1 \in \mathbb{F} = M_0$ generates M. Therefore, M is a finitely generated graded module over the ring $\mathrm{Poly}(\mathbb{F}^n)$. By Theorem 1.11, $\mathrm{Dim}\, M_D = h_M(D)$ for large D, where h_M is a polynomial. But then $\mathrm{Dim}\, R_D = \sum_{d=0}^{D} \mathrm{Dim}\, M_d$ is also a polynomial for large D.

Exercise 6.2. (The image of a polynomial map) Suppose that $P : \mathbb{F} \to \mathbb{F}^2$ is a polynomial map with coordinates P_1, P_2. Prove that the image of P lies in $Z(Q)$ for some polynomial $Q \in \mathrm{Poly}(\mathbb{F}^2)$. If P_1, P_2 have degree at most D, give an estimate for $\mathrm{Deg}\, Q$.

Here is a small generalization. Suppose that $H \subset \mathbb{F}^2$ is the hyperbola defined by the equation $x^2 - y^2 = 1$. Suppose that $P : \mathbb{F}^2 \to \mathbb{F}^2$ is a polynomial map with coordinates P_1, P_2. Prove that the image of P lies in $Z(Q)$ for some polynomial $Q \in \mathrm{Poly}(\mathbb{F}^2)$. If P_1, P_2 have degree at most D, give an estimate for $\mathrm{Deg}\, Q$.

What is the most general version of this result that you can prove?

https://doi.org/10.1090//ulect/064/07

CHAPTER 7

Incidence geometry

The main subject of this book is applying polynomial methods to incidence geometry. In the first part of the book, we introduced the polynomial method and saw some applications in a few different directions. We are now starting the second part of the book, where we introduce incidence geometry. Over the next three chapters, we will explain some of the main results, techniques, and open problems in the field. In the third part of the book, we will apply the polynomial method to incidence geometry.

Incidence geometry studies the possible intersection patterns of simple geometric objects such as lines or circles. The joints problem and the distinct distance problem are examples of incidence geometry. In the next few chapters, we consider incidence geometry more systematically. We discuss some fundamental results and questions in the field that help put these problems in context.

In this chapter, we discuss incidence geometry in the plane. We begin by studying lines in the plane and learn the Szemerédi-Trotter theorem, the most fundamental result in the field. Here we meet one of the most important discoveries in the field: the role of topology in incidence geometry. Next we consider distance problems in the plane, such as the distinct distance problem. This leads into the incidence geometry of circles and other curves. Here we meet many difficult open problems, and we try to explain why these problems are difficult with our current methods. I hope that this chapter will help to put the distinct distance problem in context.

In the next chapter we will discuss incidence geometry in three or higher dimensions. We will see some new issues that appear in higher dimensions. The joints problem is an example of incidence geometry in higher dimensions, and I hope that this chapter will help to put it in context.

In the third chapter, we discuss the method of partial symmetries, which gives an interesting connection between some planar problems - like the distinct distance problem - and problems in higher dimensions.

These three chapters give a brief introduction to incidence geometry. We mostly focus on the issues that will be relevant in the applications of the polynomial method later. But there is a lot more to the field. The book *Combinatorial geometry and its algorithmic applications* [**PS**] by Pach and Sharir gives a fuller introduction to the subject, with many more questions, results, and applications. In particular, there are interesting connections between incidence geometry and algorithms for geometric problems which are described in [**PS**].

7.1. The Szemerédi-Trotter theorem

Let $\mathfrak{L}$ denote a set of L lines in the plane $\mathbb{R}^2$. An r-rich point of $\mathfrak{L}$ is a point that lies in at least r lines. The set of all r-rich points of $\mathfrak{L}$ is denoted $P_r(\mathfrak{L})$.

$$P_r(\mathfrak{L}) := \{x \in \mathbb{R}^2 | x \text{ lies in at least } r \text{ lines of } \mathfrak{L}\}.$$

One of the basic questions about the intersection patterns of lines in the plane is to estimate

$$\max_{|\mathfrak{L}|=L} |P_r(\mathfrak{L})|.$$

This question was answered up to a constant factor by Szemerédi and Trotter in [**SzTr**] in the early 1980's. Before we state their theorem, let us consider some examples.

If we pick Lr^{-1} points, and we draw r lines through each point, we get a configuration of L lines with Lr^{-1} r-rich points. We call this the 'stars' configuration.

If we pick L generic lines, then we get $\binom{L}{2} \sim L^2$ 2-rich points, but no 3-rich points.

To construct a set of L lines with $\sim L^2$ 3-rich points, we can use a grid of vertical, horizontal, and diagonal lines as follows:

- Horizontal lines $y = b$ for each integer $b = 1, ..., L/4$.
- Vertical lines $x = a$ for each integer $a = 1, ..., L/4$.
- Diagonal lines $x - y = c$ for each integer $c = -L/4, ..., L/4 - 1$.

Now each integer point (a, b) with $1 \leq a, b \leq L/4$ is a 3-rich point, giving $\sim L^2$ 3-rich points.

For larger r, we can make many r-rich points by using a similar grid structure and adding lines of other slopes. Begin with an $N \times N$ square grid of points. Then we choose r slopes for lines. We want the slopes to be rational numbers with small numerator and denominator. Let us define the set of slopes to be the first r numbers in the list $0, 1, 1/2, 1/3, 2/3, 1/4, 3/4, 1/5, 2/5, ...$ In this list we only include fractions in lowest terms. Next we let $\mathfrak{L}$ be the set of lines that go through at least one point of the $N \times N$ grid with slope in the list of r slopes. We call these examples grid examples. We will analyze the grid examples below and see that a grid example with L lines can determine $\sim L^2 r^{-3}$ r-rich points.

The Szemerédi-Trotter theorem says that these examples are sharp up to a constant factor.

THEOREM 7.1. If $\mathfrak{L}$ is a set of L lines in the plane, then

$$|P_r(\mathfrak{L})| \lesssim L^2 r^{-3} + L r^{-1}.$$

If $r \gtrsim L^{1/2}$, then the second term dominates, and the stars example is sharp. If $r \lesssim L^{1/2}$, then the first term dominates, and the grid example is sharp.

Now we analyze the grid examples. Let S_r be the first r rational numbers in the sequence $0, 1, 1/2, 1/3, 2/3, 1/4, 3/4$, etc. Let G_N be the $N \times N$ integer grid $\{(a, b) \in \mathbb{Z}^2 | 1 \leq a, b \leq N\}$. Let $\mathfrak{L}_{N,r}$ be the set of lines with slope in S_r that intersect at least one point of G_N. By construction, every point of G_N is an r-rich point of $\mathfrak{L}_{N,r}$, and so $|P_r(\mathfrak{L}_{N,r})| \geq N^2$. It is not as obvious how big $|\mathfrak{L}_{N,r}|$ is. We will show that

$$L := |\mathfrak{L}_{N,r}| \lesssim N r^{3/2}.$$

This implies that

$$|P_r(\mathfrak{L}_{N,r})| \geq N^2 \gtrsim L^2 r^{-3}.$$

(With a little more work, the reader can show that $L \sim Nr^{3/2}$ and that $|P_r(\mathfrak{L}_{N,r})| \sim L^2 r^{-3}$.)

We claim that the r slopes in S_r all have numerator and denominator $\lesssim r^{1/2}$. If we consider all fractions p/q with $1 \le p < q \le Q$, then we have approximately $(1/2)Q^2$ fractions. Not all these fractions are in lowest terms, so they don't give $(1/2)Q^2$ rational numbers, but the number of fractions in lowest terms is still $\gtrsim Q^2$. See Exercise 7.1 for a quick proof. (In fact, something much sharper is known: the probability that a fraction from the list is in lowest terms tends to $6/\pi^2$ as $Q \to \infty$.) This implies that the fractions in S_r have numerator and denominator $\lesssim r^{1/2}$.

If l is a line with slope p/q going through an integer point $x = (x_1, x_2)$, then $(x_1 + q, x_2 + p)$ is another integer point on the line l. Since each of our slopes has numerator and denominator $\lesssim r^{1/2}$, each line in $\mathfrak{L}_{N,r}$ contains $\gtrsim Nr^{-1/2}$ points of a $3N \times 3N$ grid centered on our original $N \times N$ grid. We will assume from now on that $N \ge r^{1/2}$, so that $Nr^{-1/2} \ge 1$. (If $N < r^{1/2}$, then a typical line in $\mathfrak{L}_{N,r}$ intersects G_N in just a single point.)

Now we can estimate the number of lines in $\mathfrak{L}_{N,r}$ by a double counting argument. Each point in the $3N \times 3N$ grid lies in at most r lines of $\mathfrak{L}_{N,r}$. Since each line of $\mathfrak{L}_{N,r}$ contains $\gtrsim Nr^{-1/2}$ points of this $3N \times 3N$ grid, we see that

$$L \lesssim \frac{N^2 r}{Nr^{-1/2}} = Nr^{3/2}.$$

This finishes our proof that $|P_r(\mathfrak{L}_{N,r})| \gtrsim |\mathfrak{L}_{N,r}|^2 r^{-3}$. We have assumed that $N \ge r^{1/2}$. Since $L \sim Nr^{3/2}$, this corresponds to $L \gtrsim r^2$. If we pick L and r with $L \lesssim r^2$, then the stars example gives $\sim Lr^{-1}$ r-rich points, and if we pick L and r with $L \gtrsim r^2$, then a grid example gives $\sim L^2 r^{-3}$ r-rich points. Theorem 7.1 says that these examples are sharp up to a constant factor.

Having seen some examples, we begin to discuss upper bounds for $|P_r(\mathfrak{L})|$. The first upper bounds in this problem are based on the Euclidean axiom:

Two lines intersect in at most one point. (E)

We can use this lemma to bound $P_r(L)$ by a double counting argument:

LEMMA 7.2. $P_r(L) \le \binom{L}{2}\binom{r}{2}^{-1} \sim L^2 r^{-2}$.

PROOF. Suppose $\mathfrak{L}$ is a set of L lines. For each point $x \in P_r(\mathfrak{L})$, list all the *pairs* of lines in $\mathfrak{L}$ that intersect at x. For each x, we have a list of at least $\binom{r}{2}$ pairs. By the Euclidean axiom, any pair of lines intersects in at most one point, so the total number of pairs of lines in all these lists is at most $\binom{L}{2}$. Therefore, $|P_r(\mathfrak{L})|\binom{r}{2} \le \binom{L}{2}$. □

Is this the only bound for $P_r(L)$ that follow from the Euclidean axiom? Perhaps surprisingly, there is a subtler counting argument that gives additional bounds.

LEMMA 7.3. If $r \ge 2L^{1/2}$, then $P_r(L) < 2Lr^{-1}$.

PROOF. Suppose that $\mathfrak{L}$ is a set of L lines. We will give a proof by contradiction, so suppose that $|P_r(\mathfrak{L})| \ge 2Lr^{-1}$. Now choose a subset $P' \subset P_r(\mathfrak{L})$ with $2Lr^{-1} \le |P'| < 2Lr^{-1}$. Since $r \ge 2L^{1/2}$, $|P'| < r/2 + 1$. Each point of P' lies in r lines of $\mathfrak{L}$. But because of the Euclidean axiom, less than $r/2$ of those lines can intersect any other point of P'! Therefore, the number of lines L is bigger than $|P'|(r/2)$. But $|P'|(r/2) \ge (2Lr^{-1})(r/2) = L$. This contradiction shows that $|P_r(\mathfrak{L})| < 2Lr^{-1}$. □

In the range $r \geq 2L^{1/2}$, Lemma 7.3 matches the stars example up to a constant factor. Therefore it is sharp up to a constant factor, which proves Theorem 7.1 in the range $r \geq 2L^{1/2}$. In the range $r \leq L^{1/2}$, the method of Lemma 7.3 doesn't apply, and Lemma 7.2 gives $|P_r(\mathfrak{L})| \lesssim L^2 r^{-2}$. There is a gap between these estimates and Theorem 7.1. The gap is largest and most interesting when $r = L^{1/2}$. In this case, the counting lemmas give $|P_{L^{1/2}}(\mathfrak{L})| \lesssim L$, and Theorem 7.1 gives $|P_{L^{1/2}}(\mathfrak{L})| \lesssim L^{1/2}$.

These two lemmas give some bounds, but they don't prove Theorem 7.1. One may wonder if there is an even more clever way to use the Euclidean axiom to get better bounds. It turns out that the Euclidean axiom alone is not enough to prove Theorem 7.1. At this point in the story, we come to the main obstacle in understanding Theorem 7.1. The Euclidean axiom holds over every field, but the Szemerédi-Trotter theorem does not hold over every field.

Let $\mathbb{F}_q$ be the finite field with q elements. Let $\mathfrak{L}$ be the set of all non-vertical lines in $\mathbb{F}_q^2$. Each non-vertical line is a graph $y = mx + b$, so there are exactly $L = q^2$ lines in $\mathfrak{L}$. Each point of $\mathbb{F}_q^2$ lies in one line of every slope, so there are q^2 points in $P_q(\mathfrak{L})$. So we have $|P_{L^{1/2}}(\mathfrak{L})| = L$, essentially matching the upper bound from the counting lemmas. (If we include vertical lines, the example becomes even slightly stronger – see Exercise 7.3.)

We have now seen the main obstacle in proving Theorem 7.1. In order to prove Theorem 7.1, we need to use something that is true for lines in $\mathbb{R}^2$ but false for lines in $\mathbb{F}_q^2$. We can't just use the fact that two lines intersect in at most one point. We need to use something rather different and subtler.

There are several proofs of Theorem 7.1. All of the proofs use the topology of $\mathbb{R}^2$ in some way. Understanding how to use topology to bound quantities like $|P_r(\mathfrak{L})|$ is one of the main discoveries of incidence geometry. Over the course of the book, we will give two proofs of Theorem 7.1 and sketch a third. In this chapter, we will give a proof that uses Euler's formula from topology.

7.1.1. Exercises.

Exercise 7.1. Let $R(Q)$ be the set of pairs (p, q) with $1 \leq p < q \leq Q$ and with $\gcd(p, q) = 1$. Prove that $|R(Q)| \geq (1/10)Q^2$. We sketch one possible proof below.

For an integer $a \geq 2$, let $D(a, Q)$ be the set of pairs (p, q) with $1 \leq p < q \leq Q$, and so that a divides both p and q. Check that $|D(a, Q)| \leq a^{-2}Q^2$.

Note that $R(Q)$ can be formed by starting with all pairs (p, q) with $1 \leq p < q \leq Q$ and then removing $D(a, Q)$ for every prime a. Therefore,

$$|R(Q)| \geq Q^2 - \sum_{a \text{ prime}} |D(a, Q)| \geq \left(1 - \sum_{a \geq 2, \text{ prime}} a^{-2}\right) Q^2.$$

So it only remains to check that the quantity in parentheses is at least $1/10$. To do this, consider the sum in each dyadic interval:

$$\sum_{2^k \leq a < 2^{k+1}, a \text{ prime}} a^{-2} \leq \sum_{2^k \leq a < 2^{k+1}} a^{-2} < 2^{-k}.$$

Using this crude bound for $k \geq 2$ to bound all the terms $a \geq 4$ and estimating $a = 2, 3$ by hand, prove that $|R(Q)| \geq (1/10)Q^2$.

Exercise 7.2. Show that $|\mathfrak{L}_{N,r}| \sim Nr^{3/2}$. We proved in the text that $|\mathfrak{L}_{N,r}| \lesssim Nr^{3/2}$, and so it just remains to prove that $|\mathfrak{L}_{N,r}| \gtrsim Nr^{3/2}$. Hint: Let $S_r' \subset S_r$

be the subset of fractions with numerator and denominator at least $\frac{1}{10}r^{1/2}$. Check that $|S'(r)| \geq r/2$. Let $\mathfrak{L}'_{N,r} \subset \mathfrak{L}_{N,r}$ be the lines with slope in S'_r. Each line in $\mathfrak{L}'_{N,r}$ intersects the $N \times N$ grid G_N in $\lesssim Nr^{-1/2}$ points. But each point of G_N lies in at least $r/2$ lines of $\mathfrak{L}'_{N,r}$. Use double counting to finish the argument.

EXERCISE 7.3. There is a slighlty sharper version of the finite field example that we mentioned above, using the projective plane $\mathbb{PF}_q^2$ instead of the regular plane $\mathbb{F}_q^2$. Let $\mathfrak{L}$ denote the set of all lines in the projective space $\mathbb{PF}_q^2$. Check that every two lines of $\mathfrak{L}$ intersect at exactly one point in $\mathbb{PF}_q^2$ and each point of $\mathbb{PF}_q^2$ lies in exactly $q+1$ lines of $\mathfrak{L}$. Therefore, we get equality in Lemma 7.2 with $r = q+1$:

$$|P_{q+1}(\mathfrak{L})| = \binom{|\mathfrak{L}|}{2} / \binom{q+1}{2}.$$

7.2. Crossing numbers and the Szemerédi-Trotter theorem

In this section, we give a proof of the Szemerédi-Trotter theorem using the Euler formula from topology. To see how the Euler formula may come into play, let us first quickly prove something much weaker. Is it possible to arrange L lines in the plane so that $|P_r(\mathfrak{L})|$ is exactly $\binom{L}{2}\binom{r}{2}^{-1}$, where $L > r$? For this to happen, every pair of lines must intersect, and every intersection point must lie in exactly r lines. If $r-1$ is a prime power, then we saw in Exercise 7.3 that this can happen for lines in $\mathbb{PF}_{r-1}^2$. But we claim that for $r \geq 6$, this is impossible in $\mathbb{R}^2$ because of Euler's formula. A set $\mathfrak{L}$ of lines in $\mathbb{R}^2$ determines a polyhedral structure on $\mathbb{R}^2$, where each intersection point is a vertex, each segment of a line between two vertices is an edge, and each component of the complement of the lines is a face. Some of the faces and edges are unbounded. We let X be the union of the bounded faces, edges, and vertices. Since $L > r$, X is non-empty, and it must be a topological disk. The lines define a polyhedral structure on X with V vertices, E edges, and F faces. By Euler's formula, we have $V - E + F = 1$. Each vertex has degree at least r, and each edge contains only two vertices. Therefore, $V \leq (2/r)E \leq (1/3)E$. Each face has at least three edges, and each edge lies in at most two faces, and so $F \leq (2/3)E$. But then the left-hand side, $V - E + F$, is at most 0, giving a contradiction.

Directly generalizing the argument in the last paragraph doesn't give good estimates for $|P_r(\mathfrak{L})|$. We encourage the reader to try this and see what happens. Crossing numbers of graphs will allow us to leverage the Euler formula more effectively by looking at a well-chosen piece of the above picture instead of the whole picture. This approach to proving the Szemerédi-Trotter theorem is due to Székely, [**Sz**]. It is based on an important estimate about crossing numbers of graphs, proven independently by Leighton ([**Le**]) and Ajtai, Chvátal, Newborn and Szemerédi ([**ACNS**]).

7.2.1. Crossing number estimates. A drawing of a graph G in the plane assigns each vertex of G to a distinct point in $\mathbb{R}^2$ and assigns each edge of G to a simple continuous curve in $\mathbb{R}^2$ between its endpoints so that no edge passes through any vertex other than its endpoints. Every finite graph admits a drawing in the plane: map the vertices to points in general position and map the edges to straight lines. A crossing in the drawing is a point x and an unordered pair of open edges e, e', so that $x \in e \cap e'$. For example, if r distinct edges go through a point, then it counts as $\binom{r}{2}$ crossings. The crossing number of G is the smallest number of

crossings among all drawings of G in the plane. The graph G is planar if and only if the crossing number of G is zero. We denote the crossing number of G by $k(G)$.

For example, consider the complete graph on N vertices, K_N. How does the crossing number of K_N grow as $N \to \infty$? The graph K_N has $\binom{N}{2}$ edges. If a graph has E edges, then we can draw it with straight line edges giving at most $\binom{E}{2}$ crossings, and so the crossing number of K_N is clearly at most $\binom{\binom{N}{2}}{2} \sim N^4$. We will prove that the crossing number of K_N really does grow like N^4. More generally, we will study how the crossing number of a graph G is related to the number of edges and vertices of G.

It's also interesting to consider crossings in straight line drawings. We let $k_{str}(G)$ denote the straight-line crossing number of G: the minimal number of crossings in a drawing of G where each edge is a straight line. Clearly $k_{str}(G) \geq k(G)$. This straight-line crossing number $k_{str}(G)$ may be different from $k(G)$. Either one is useful for proving the Szemerédi-Trotter theorem. The straight-line crossing number requires a little less topology background to work with, but $k(G)$ will have later applications to incidence geometry problems with curves.

Now we turn to estimates about the crossing numbers of graphs, starting with a classical estimate for the number of edges and vertices of a planar graph. This combinatorial estimate is based on the Euler formula.

PROPOSITION 7.4. If G is a planar graph with E edges and V vertices, then $E - 3V \leq -6$.

We first sketch the main idea of the proof. Suppose that G is planar and consider an embedding of G into S^2. This embedding cuts S^2 into faces, and we get a polyhedral structure on S^2 with V vertices, E edges, and some number F of faces. By the Euler formula, $V - E + F = 2$. The number of faces cannot easily be read from the graph G, but we can estimate it as follows. Each face has at least three edges in its boundary, whereas each edge borders exactly two faces. Therefore $F \leq (2/3)E$. Plugging in we get $2 = V - E + F \leq V - (1/3)E$. Rearranging gives $E - 3V \leq -6$.

This sketch is not quite a proof. For example, suppose that the graph G is a disconnected graph homeomorphic to two circles. If G is drawn in S^2 as two concentric circles, then there are three "faces" : two disks and an annulus. But the Euler formula is false for this configuration, because annular faces are not allowed. The Euler formula also does not apply if G consists of a single edge, or if G is a tree. To apply the Euler formula, we need to know that each face is homeomorphic to a polygon.

Writing down all the details of the argument is actually a little bit involved, although the main idea is elegant. When I was teaching this proof, it reminded me of Lakatos's book [**Lak**] about the history of the Euler formula and the long story of clarifying the statement of the result.

We invite the interested reader to finish the proof as an exercise, and we give a few more clues for guidance. Alternatively, the reader may consult the original papers [**Le**] or [**ACNS**]. Suppose that G is a connected planar graph and that every vertex belongs to at least two edges. If we draw G in S^2, then we claim that each component of $S^2 \setminus G$ will be homeomorphic to a polygon, and the Euler formula argument above will apply. On the other hand, once we know that the result holds

for connected planar graphs where every vertex belongs to at least two edges, then it follows easily for all planar graphs by induction on the number of vertices.

As a consequence of Proposition 7.4, it follows that the graph K_5 is not planar. The complete graph K_5 has 5 vertices and 10 edges, and so $E - 3V = 10 - 15 = -5 > -6$. It's not hard to draw K_5 with one crossing, and so we see that $k(K_5) = 1$.

If $E - 3V > -6$, then we see that G is not planar, and if $E - 3V$ is large and positive then we may expect that G has a large crossing number. We now prove a simple bound of this type.

PROPOSITION 7.5. The crossing number of G is at least $E - 3V + 6$.

PROOF. Let $k(G)$ be the crossing number of G. Embed G in the plane with $k(G)$ crossings. By removing at most $k(G)$ edges, we get a planar graph G' with $E' = E - k$ edges and $V' \leq V$ vertices. By Proposition 7.4 we see that $-6 \geq E' - 3V' \geq E - k - 3V$. □

For the complete graph K_N, this Proposition gives $k(K_N) \geq \binom{N}{2} - 3N + 6 \sim N^2$. We will eventually prove a much stronger estimate: $k(K_N) \gtrsim N^4$.

How can we hope to improve Proposition 7.5? When we remove an edge of G, it's in our interest to remove the edge with the most crossings, and when we do this, the crossing number of G can decrease by more than 1. For example, for the complete graph K_N, it looks plausible that there is always an edge with $\sim N^2$ crossings. But how can we prove such an estimate?

This seems to be a tricky problem, and [**Le**] and [**ACNS**] found a clever solution. Instead of trying to prove that one edge of G intersects many other edges of G, we consider a small random subgraph $G' \subset G$ and prove that some edges of G' must cross. Since G' is only a small piece of G, it follows that many pairs of edges in G must cross.

THEOREM 7.6. ([**Le**] and [**ACNS**]) If G is a graph with E edges and V vertices, and $E \geq 4V$, then the crossing number of G is at least $(1/64)E^3V^{-2}$.

In particular, for the complete graph K_N, Theorem 7.6 implies that $k(K_N) \gtrsim \binom{N}{2}^3 N^{-2} \sim N^4$.

We start with a slightly easier estimate for the straight-line crossing number and then explain the modifications needed for the general crossing number.

THEOREM 7.7. ([**Le**] and [**ACNS**]) If G is a graph with E edges and V vertices, and $E \geq 4V$, then $k_{str}(G) \geq (1/64)E^3V^{-2}$.

PROOF. Let p be a number between 0 and 1 which we choose below. Let G' be a random subgraph of G formed by including each vertex of G independently with probability p. We include an edge of G in G' if its endpoints are in G'.

We consider the expected values for the number of vertices and edges in G'. The expected value of V' is pV. The expected value of E' is p^2E. For every subgraph $G' \subset G$, the crossing number of G' is at least $E' - 3V'$. Therefore, the expected value $k(G')$ is at least $p^2E - 3pV$. Since the straight line crossing number is at least the crossing number, the expected value of $k_{str}(G')$ is at least $p^2E - 3pV$.

On the other hand, we give an upper bound on the expected value of $k_{str}(G')$ as follows. Consider a straight-line drawing of G in $\mathbb{R}^2$ with $k_{str}(G)$ crossings. Because each edge is a straight line, two edges sharing a common vertex can never cross. So each crossing must involve two edges containing four distinct vertices.

By restricting our drawing of G to G', we get a straight-line drawing of G'. The expected number of crossings in this drawing of G' is $p^4 k_{str}(G)$. Therefore, the expected value of $k_{str}(G')$ is at most $p^4 k_{str}(G)$.

Comparing our upper and lower bounds for the expected value of $k_{str}(G')$, we see that $p^4 k_{str}(G) \geq p^2 E - 3pV$, and so we get the following lower bound for $k_{str}(G)$.

$$k_{str}(G) \geq p^{-2} E - 3p^{-3} V.$$

We can now choose p to optimize the right-hand side. We choose $p = 4V/E$, and we have $p \leq 1$ since we assumed $4V \leq E$. Plugging in we get $k(G) \geq (1/64) E^3 V^{-2}$. □

Next we discuss the modifications needed to prove Theorem 7.6. The new issue is that in a drawing of G with curved edges, two edges leaving the same vertex may cross. Such a crossing involves two edges with a total of three vertices. So this crossing would appear in G' with probability p^3 much higher than p^4. We get around this difficulty by proving the following lemma.

LEMMA 7.8. In a drawing of a graph G with $k(G)$ crossings, each crossing involves two edges with four distinct vertices.

We describe the idea of the proof of this lemma. Suppose that we have a drawing of a graph G where two edges sharing a vertex have a crossing. Given a drawing with such a crossing, we explain how to modify it to reduce the crossing number. This will show that in a drawing with the minimal number of crossings, two edges leaving a common vertex cannot cross.

Suppose that e_1 and e_2 each leave the vertex v and cross at a point x. (If they cross several times, then let x be the last crossing.) We modify the drawing as follows. Suppose that e_1 crosses k_1 other edges on the way from v to x and that e_2 crosses k_2 other edges on the way from v to x. We label the edges so that so that $k_1 \leq k_2$. Then we modify the drawing of e_2 so that e_2 follows parallel to e_1 until x and then rejoins its original course at x, so that e_1 and e_2 never cross. This operation reduces the number of crossings in the drawing.

It takes a little work to make this argument rigorous. We refer to the original papers [**Le**] and [**ACNS**] or leave it as an exercise for the reader. The argument is a little bit easier if we assume that the edges are piecewise smooth curves instead of just continuous curves. One can define the piecewise smooth crossing number using piecewise smooth embeddings and prove a piecewise smooth crossing number estimate. Such an estimate will be sufficient for all the applications that we talk about later.

7.2.2. The Szemerédi-Trotter theorem. Now we use our crossing number estimate to prove the Szemerédi-Trotter theorem.

THEOREM 7.9. If $\mathfrak{L}$ is a set of L lines in the plane, then

$$|P_r(\mathfrak{L})| \leq \max(2Lr^{-1}, 2^9 L^2 r^{-3}).$$

PROOF. Using $\mathfrak{L}$ and $P_r(\mathfrak{L})$ we construct a graph drawn in the plane. The vertices of our graph G are the points of $P_r(\mathfrak{L})$. We join two vertices with an edge of G if the two points are two consecutive points of $P_r(\mathfrak{L})$ along a line $l \in \mathfrak{L}$. The crossing number of this drawing is at most $\binom{L}{2} \leq L^2$, since each crossing of the

graph G must correspond to an intersection of two lines of $\mathfrak{L}$. This is also a straight line drawing. So we see that $k(G) \leq k_{str}(G) \leq L^2$.

We will count the vertices and edges of the graph G and apply the crossing number theorem. The number of vertices of our graph is $V = |P_r(\mathfrak{L})|$. The number of edges of our graph is $rV - L$. To see this, we count the number of pairs (v, e) where v is a vertex and e is an edge containing v. Each vertex is adjacent to $2r$ segments of lines. The bounded segments are edges, but there are also two unbounded segments on each line. Therefore, there are $2rV - 2L$ pairs (v, e) as above. Each edge contains exactly two vertices, so the number of edges is $rV - L$. As long as $E \geq 4V$, we can apply the (straight line) crossing number theorem, Theorem 7.7, and it gives

$$L^2 \geq k_{str}(G) \geq (1/64)(rV - L)^3 V^{-2}.$$

The rest of the argument is just elementary computation. The number of edges is $rV - L$. The most interesting case occurs when the first term dominates. Suppose that $rV - L \geq (1/2)rV$. In this case, we have $L^2 \geq 2^{-9} r^3 V$, and so $|P_r(\mathfrak{L})| = V \leq 2^9 L^2 r^{-3}$. On the other hand, if $rV - L < (1/2)rV$, then we have $|P_r(\mathfrak{L})| = V \leq 2Lr^{-1}$.

In order to apply the crossing number theorem, we assumed that $E \geq 4V$. If $E < 4V$, we have $rV - L \leq 4V$, and hence $V \leq \frac{L}{r-4}$. As long as $r \geq 8$, this implies $V \leq 2L/r$, and we are done. Finally, for $r < 8$, the counting bound $|P_r(\mathfrak{L})| \leq \binom{L}{2}/\binom{r}{2}$ does the job: $\binom{L}{2}/\binom{r}{2} \leq 2L^2 r^{-2} \leq 2^9 L^2 r^{-3}$. □

Exercise 7.4. In this exercise, we describe an alternate example showing that the Szemerédi-Trotter theorem is sharp. This example uses a rectangular grid instead of a square grid.

Let $\mathfrak{L}$ be the set of lines in the plane of the form $y = mx + b$, where m, b are integers in the ranges $1 \leq m \leq r$ and $|b| \leq 2H$. Here r and H are parameters which can be any positive integers. If $(x, y) \in \mathbb{Z}^2$ with $|x| \leq Hr^{-1}$ and $|y| \leq H$, prove that (x, y) is an r-rich point for $\mathfrak{L}$. Conclude that

$$|P_r(\mathfrak{L})| \gtrsim L^2 r^{-3}.$$

7.3. The language of incidences

The Szemerédi-Trotter theorem plays a fundamental role in incidence geometry in the plane. There are several closely related versions of the Szemerédi-Trotter estimate, and each version is valuable. Theorem 7.1 gives an estimate for the number of r-rich points of a set of lines. Next we introduce the definition of an incidence and use it to give a different formulation of the theorem.

If $\mathcal{S}$ is a set of points and $\mathfrak{L}$ is a set of lines (or curves), the set of incidences is defined as follows:

$$I(\mathcal{S}, \mathfrak{L}) = \{(p, l) \in \mathcal{S} \times \mathfrak{L} | p \in l\}.$$

The number of incidences can be counted in several ways, making it a useful object in double counting arguments. For instance,

$$|I(\mathcal{S}, \mathfrak{L})| = \sum_{p \in \mathcal{S}} \big|\{l \in \mathfrak{L} | p \in l\}\big| = \sum_{l \in \mathfrak{L}} \big|\{p \in \mathcal{S} | p \in l\}\big|.$$

Here is a basic estimate for $|I(\mathcal{S}, \mathfrak{L})|$ that takes advantage of the multiple ways to count $I(\mathcal{S}, \mathfrak{L})$.

PROPOSITION 7.10. If $\mathcal{S}$ is a set of S points in the plane, and $\mathfrak{L}$ is a set of L lines in the plane, then

$$|I(\mathcal{S}, \mathfrak{L})| \leq L + S^2,$$

and

$$|I(\mathcal{S}, \mathfrak{L})| \leq S + L^2.$$

PROOF. We let $\mathfrak{L}_1$ be the set of lines in $\mathfrak{L}$ that contain exactly one point of $\mathcal{S}$, and we let $\mathfrak{L}_{>1}$ be the set of lines in $\mathfrak{L}$ that contain more than one point of $\mathcal{S}$. We see that $|I(\mathcal{S}, \mathfrak{L})| = |I(\mathcal{S}, \mathfrak{L}_1)| + |I(\mathcal{S}, \mathfrak{L}_{>1})|$.

Now $|I(\mathcal{S}, \mathfrak{L}_1)| \leq |\mathfrak{L}_1| \leq L$. On the other hand,

$$|I(\mathcal{S}, \mathfrak{L}_{>1})| = \sum_{p \in \mathcal{S}} |\{l \in \mathfrak{L}_{>1} | p \in l\}|.$$

For each $p \in \mathcal{S}$, there are at most $S - 1$ lines containing p that hit another point of $\mathcal{S}$. Therefore,

$$|I(\mathcal{S}, \mathfrak{L}_{>1})| \leq \sum_{p \in \mathcal{S}} (S - 1) \leq S^2.$$

This finishes the proof of the first inequality. The second inequality is similar. We define $\mathcal{S}_1$ to be the set of points of $\mathcal{S}$ that lie in exactly one line of $\mathfrak{L}$ and $\mathcal{S}_{>1}$ to be the set of points of $\mathcal{S}$ that lie in more than one line of $\mathfrak{L}$.

$$|I(\mathcal{S}, \mathfrak{L})| \leq |I(\mathcal{S}_1, \mathfrak{L})| + |I(\mathcal{S}_{>1}, \mathfrak{L})| \leq S + \sum_{l \in \mathfrak{L}} |\{p \in \mathcal{S}_{>1} | p \in l\}|.$$

For each $l \in \mathfrak{L}$, there are at most $L - 1$ points in l that intersect another line of $\mathfrak{L}$, and so the last expression is

$$\leq S + \sum_{l \in \mathfrak{L}} (L - 1) \leq S + L^2.$$

□

We now state the Szemerédi-Trotter bound on the number of incidences.

THEOREM 7.11. If $\mathcal{S}$ is a set of S points in the plane, and $\mathfrak{L}$ is a set of L lines in the plane, then the number of incidences between $\mathcal{S}$ and $\mathfrak{L}$ is bounded as follows:

$$|I(\mathcal{S}, \mathfrak{L})| \lesssim (S^{2/3} L^{2/3} + S + L).$$

Szemerédi and Trotter proved Theorem 7.11 and Theorem 7.1 in [**SzTr**]. The two theorems are closely related, and either estimate can be referred to as the Szemerédi-Trotter theorem. Theorem 7.11 can be proved by making a small modification in the proof of Theorem 7.1 in Subsection 7.2.2.

The two results are also essentially equivalent. We now prove Theorem 7.11 using Theorem 7.1.

PROOF. We sort $\mathcal{S}$ according to the number of lines that each point of $\mathcal{S}$ lies in. We define $\mathcal{S}_{high}$ to be the set of points $p \in \mathcal{S}$ that lie in at least $2L^{1/2}$ lines of $\mathfrak{L}$. By Proposition 7.10, $|I(\mathcal{S}_{high}, \mathfrak{L})| \leq |\mathcal{S}_{high}|^2 + L$. But $|\mathcal{S}_{high}| \leq |P_{2L^{1/2}}(\mathfrak{L})|$. By Lemma 7.3, $|P_{2L^{1/2}}(\mathfrak{L})| \leq L^{1/2}$. All together, we get $|I(\mathcal{S}_{high}, \mathfrak{L})| \leq 2L$. This inequality is based only on double counting and not on topological considerations.

We subdivide the rest of $\mathcal{S}$ more finely. We define

$$\mathcal{S}_k := \{p \in \mathcal{S} \text{ so that } 2^{k-1} \leq |\{l \in \mathfrak{L} | p \in l\}| < 2^k\}.$$

From this definition, we see that $|I(\mathcal{S}_k, \mathfrak{L})| \leq 2^k |S_k|$. We let K denote the smallest integer with $2^{K-1} \geq 2L^{1/2}$. This guarantees that $\cup_{k>K} \mathcal{S}_k \subset \mathcal{S}_{high}$. We can now break up the incidences as follows:

$$|I(\mathcal{S}, \mathfrak{L})| \leq 2L + \sum_{k=1}^{K} 2^k |S_k|.$$

We know that $|S_k| \leq |P_{2^{k-1}}(\mathfrak{L})|$. Since $2^{k-1} \leq L^{1/2}$, Theorem 7.1 tells us that $|P_{2^{k-1}}(\mathfrak{L})| \lesssim L^2 2^{-3k}$. Of course we also know that $|S_k| \leq S$. Putting everything together, we see that

$$|I(\mathcal{S}, \mathfrak{L})| \lesssim L + \sum_{k=1}^{K} 2^k \min(S, 2^{-3k} L^2).$$

It just remains to estimate this sum. The sum splits into two pieces, depending on whether S or $2^{-3k}L^2$ is bigger, and each piece is just a geometric series. We note that $S \leq 2^{-3k}L^2$ if and only if $2^k \leq L^{2/3} S^{-1/3}$. So the last sum is bounded by

$$\sum_{1 \leq k \leq K; 2^k \leq L^{2/3} S^{-1/3}} 2^k S + \sum_{1 \leq k \leq K; 2^k \geq L^{2/3} S^{-1/3}} 2^{-2k} L^2.$$

The most interesting case is when $1 \leq L^{2/3} S^{-1/3} \leq 2^K$. In this case, each geometric sum is $\lesssim L^{2/3} S^{2/3}$.

If $L^{2/3} S^{-1/3} > 2^K$, then only the first sum appears, and the total is $\lesssim 2^K S \leq L^{2/3} S^{-1/3} S = L^{2/3} S^{2/3}$.

If $L^{2/3} S^{-1/3} < 1$, then only the second sum appears, and the total is $\lesssim L^2$. But since $L^{2/3} S^{-1/3} < 1$, we have $L^2 < S$.

Combining all the cases, we see that $|I(\mathcal{S}, \mathfrak{L})| \lesssim L^{2/3} S^{2/3} + S + L$ as desired. □

Exercise 7.5. Show that Theorem 7.11 implies Thoerem 7.1.

Exercise 7.6. Prove Theorem 7.11 directly using the crossing number method, modifying the proof from Subsection 7.2.2.

Exercise 7.7. Here is a third version of the Szemerédi-Trotter estimate. Suppose that $\mathcal{S}$ is a set of S points in the plane, and let $\mathfrak{L}_r$ be the set of lines that contain $\geq r$ points of $\mathcal{S}$.

Using Theorem 7.11 or Theorem 7.1, prove that $|\mathfrak{L}_r| \leq C(S^2 r^{-3} + S r^{-1})$.

The reader may note that in the Szemerédi-Trotter theorem, the role of points and lines is symmetric: in Theorem 7.11, the right-hand side, $(S^{2/3} L^{2/3} + S + L)$, is symmetric with respect to S and L. This symmetry comes from the notion of duality in the projective plane, which interchanges points and lines. To end this section, we briefly recall how duality works in projective geometry.

If $\mathbb{F}$ is any field, then recall that the projective space $\mathbb{FP}^n$ is defined to be the set of all 1-dimensional subspaces of $\mathbb{F}^{n+1}$. We define two elements $x, y \in \mathbb{F}^{n+1} \setminus \{0\}$ to be equivalent if and only if $x = \lambda y$ for some $\lambda \in \mathbb{F}^*$. Then $\mathbb{FP}^n$ is the set of equivalence classes of elements in $\mathbb{F}^{n+1} \setminus \{0\}$. If $(x_1, x_2, ..., x_{n+1}) \in \mathbb{F}^{n+1} \setminus \{0\}$, then we write $[x_1, x_2, ..., x_{n+1}]$ for the corresponding point in $\mathbb{FP}^n$.

A k-dimensional plane in $\mathbb{FP}^n$ corresponds to a $(k+1)$-dimensional subspace $S \subset \mathbb{F}^{n+1}$. The k-plane corresponding to S is the set of all the 1-dimensional subspaces of $\mathbb{F}^{n+1}$ that are contained in S. In particular, a line in $\mathbb{FP}^2$ corresponds

to a 2-dimensional subspace of $\mathbb{F}^3$. Every 2-dimensional subspace of $\mathbb{F}^3$ is defined by an equation of the form $\sum_j a_j x_j = 0$, where $(a_1, a_2, a_3) \neq 0 \in \mathbb{F}^3$. A point $[x_1, x_2, x_3] \in \mathbb{FP}^2$ lies in this plane if and only if $\sum_j a_j x_j = 0$. Two equations $\sum j = 1^3 a_j x_j = 0$ and $\sum_{j=1}^3 b_j x_j = 0$ determine the same subspace of $\mathbb{F}^3$ if and only if $a = \lambda b$ for some $\lambda \in \mathbb{F}^*$ if and only if $[a] = [b]$. Therefore, the set of all lines in $\mathbb{FP}^2$ can be identified with $\mathbb{FP}^2$.

Given a point $p = [x_1, x_2, x_3]$ in $\mathbb{FP}^2$, we let the dual line p^* be the line corresponding to $[x_1, x_2, x_3]$. In other words, $[y] \in p^*$ if and only if $\sum_j x_j y_j = 0$. Similarly, if l is a line in $\mathbb{FP}^2$ corresponding to $[a_1, a_2, a_3]$, then we let the dual point l^* be the point $[a_1, a_2, a_3]$. Now we note that $p \in l$ if and only if $l^* \in p^*$ if and only if $\sum_j a_j x_j = 0$.

If $\mathcal{S}$ is a set of points in $\mathbb{FP}^2$ then we let $\mathcal{S}^*$ be the set of lines dual to the points of $\mathcal{S}$. Similarly, if $\mathfrak{L}$ is a set of lines in $\mathbb{FP}^2$, then we let $\mathfrak{L}^*$ be the set of points dual to the lines of $\mathfrak{L}$. The number of incidences between $\mathcal{S}$ and $\mathfrak{L}$ is the same as the number of incidences between $\mathfrak{L}^*$ and $\mathcal{S}^*$.

Finally we note that the Szemerédi-Trotter theorem extends to points and lines in $\mathbb{RP}^2$. The set of points of the form $[a_1, a_2, 1] \in \mathbb{RP}^2$ is naturally identified with $\mathbb{R}^2$. The remaining points, the points of the form $[a_1, a_2, 0]$, can be naturally identified with $\mathbb{RP}^1$, and they are called the points at infinity. For any finite set $\mathcal{S} \subset \mathbb{RP}^2$, there is a projective transformation that takes $\mathcal{S}$ into $\mathbb{R}^2 \subset \mathbb{RP}^2$. We recall the definition of a projective transformation. If $L : \mathbb{F}^{n+1} \to \mathbb{F}^{n+1}$ is an isomorphism, then L maps k-dimensional subspaces of the domain to k-dimensional subspaces of the range for every k. In particular, L induces a map from $\mathbb{FP}^n$ to $\mathbb{FP}^n$. Such a map is called a projective transformation. Since L takes each $(k+1)$-dimensional subspace of the domain to a $(k+1)$-dimensional subspace of the range, the induced map sends each k-plane in $\mathbb{FP}^n$ to another k-plane in $\mathbb{FP}^n$.

Exercise 7.8. For any finite set $\mathcal{S} \subset \mathbb{RP}^2$, prove that there is a projective transformation which maps $\mathcal{S}$ to a subset of $\mathbb{R}^2 \subset \mathbb{RP}^2$.

Prove that the Szemerédi-Trotter theorem extends to points and lines in $\mathbb{RP}^2$.

Also, using duality, show that the function

$$I_{max}(S, L) := \max_{\mathcal{S} \subset \mathbb{R}^2, |\mathcal{S}| = S, \mathfrak{L} \text{ a set of } L \text{ lines in } \mathbb{R}^2} I(\mathcal{S}, \mathfrak{L})$$

is symmetric in S, L.

Exercise 7.9. Suppose that $\mathfrak{L}$ is a finite set of lines in $\mathbb{RP}^2$. We say the lines of $\mathfrak{L}$ are concurrent if there is a single point which lies in all the lines of $\mathfrak{L}$. If the lines of $\mathfrak{L}$ are not concurrent, prove that each component of the complement of the lines is a polygon with at least three sides. Using Euler's formula, prove that there is a point lying in exactly two lines of $\mathfrak{L}$.

Using duality prove the following corollary. Suppose that $\mathcal{S}$ is a finite set of points in $\mathbb{R}^2$, not all lying on one line. Then, there is a line that contains exactly two points of $\mathcal{S}$. Such a line is called an ordinary line.

This result is called the Sylvester-Gallai theorem and the proof sketched in this exercise is due to Melchior. It plays a role in the paper [**GT**] which we will discuss in Section 7.5 below.

7.4. Distance problems in incidence geometry

Having studied one of the central theorems in incidence geometry, we now describe some other problems in the field, mostly difficult open problems. We begin with Erdős's paper "On sets of distances of n points" [**Er1**], one of the earliest papers in the field. He posed two main questions in the paper: the distinct distance problem and the unit distance problem. These questions played an important role as the field developed.

The distinct distance problem. Given N points in the plane, what is the smallest number of distinct distances they can determine? In other words, if $P \subset \mathbb{R}^2$ is a set of N points, and $d(P) := \{|p_1 - p_2|\}_{p_1,p_2 \in P, p_1 \neq p_2}$, then what is the smallest possible size of $d(P)$?

The unit distance problem. Given N points in the plane, what is the maximum number of unit distances that they can determine? In other words, if $P \subset \mathbb{R}^2$ is a set of N points, what is the maximum possible number of pairs $p_1, p_2 \in P$ with $|p_1 - p_2| = 1$?

We discuss the examples and the known bounds in these problems, starting with the distinct distance problem. A generic set of N points has $\binom{N}{2}$ distinct distances. If we arrange N points along a line with even spacing, then $|d(P)| = N - 1$. Erdős realized that it is slightly better to arrange the points in a square grid. Let P be the set of integer points (a, b) with $1 \leq a, b \leq S$, so that $N = S^2$. The distance from (a_1, b_1) to (a_2, b_2) is the square root of $(a_1 - a_2)^2 + (b_1 - b_2)^2$. We have $|(a_1 - a_2)^2| \leq S^2 = N$, and so $(a_1 - a_2)^2 + (b_1 - b_2)^2$ is an integer in the range $[1, 2N]$. Therefore, $|d(P)| \leq 2N$. But the key point is that $(a_1 - a_2)^2 + (b_1 - b_2)^2$ is a sum of two squares. Not all integers in $[1, 2N]$ are sums of two squares, and in fact most of them are not. Erdős had been studying this type of number theory problem, and using the tools in the area, he proved that $|d(P)| \sim N(\log N)^{-1/2}$. Erdős conjectured that the square grid is essentially the most extreme example, and that any set of N points has $|d(P)| \gtrsim N(\log N)^{-1/2}$. He reported that he had thought about the problem for several years, but he could only prove that $|d(P)| \gtrsim N^{1/2}$.

No one has ever found an example significantly better than the square grid. Also, as far as I know, the only examples with at most $N/2$ distinct distances come from lattices with algebraic entries - slight variations of Erdős's example.

Over the years, different people improved the estimate in the distinct distance problem. Szemerédi and Trotter were partly motivated by this problem when they proved Theorem 7.1, and ideas from that theorem helped prove some of the estimates. Using the crossing number method, Szekely proved that $|d(P)| \gtrsim N^{4/5}$, and we will study his proof in the exercises below. The best estimate before the polynomial method was $|d(P)| \gtrsim N^{.86}$ by [**KatTar**].

Using the polynomial method, the paper [**GK2**] proved that $|d(P)| \gtrsim N(\log N)^{-1}$, nearly proving Erdős's conjecture. This theorem is the main result of the book.

Next we discuss the unit distance problem. If we let P be the set of points $(1, 0), (2, 0), ..., (N, 0)$, then P has $N - 1$ unit distances. It is possible to find sets with a superlinear number of unit distances. For example, let $v_1, ..., v_a$ be unit vectors, and consider the $N = 2^a$ points $\sum_j c_j v_j$ with $c_j \in \{0, 1\}$. Each point

has a unit distances to other points, so the total number of pairs at unit distance is $a2^a = N(\log_2 N)$. Erdős considered a square grid of points with a well-chosen spacing. This example has more unit distances than the last one. Nevertheless, the known examples have $\leq C_\epsilon N^{1+\epsilon}$ unit distances.

The paper [**Er1**] proved that the number of unit distances of N points is $\lesssim N^{3/2}$. Using techniques related to Theorem 7.1, Szemerédi, Trotter, and Spencer proved in the early 80's that the number of unit distances is $\lesssim N^{4/3}$. (In the exercises, we will give a proof below using the crossing number method.) The exponent 4/3 has not been improved in thirty years, and this is a major open problem in the field.

Let me take a little time to discuss why I think these problems are interesting and important. Erdős's background was in number theory and especially the theory of prime numbers. In these areas, there are many difficult questions with elementary statements, and this feature was very interesting to Erdős. At the time, it may have seemed that deep elementary questions are a special feature of number theory. Over his career, Erdős sought out difficult elementary questions, and he found many of them. The study of prime numbers is a kind of well containing huge numbers of deep elementary questions. Erdős helped to find several new wells of deep elementary questions, and incidence geometry is one of those wells. The existence of deep elementary questions is a significant fact about mathematics, and finding new ones helps to expand our subject. In the search for difficult elementary questions, the most exciting discovery is an elementary question which looks very different from the known library of difficult elementary questions. (The most exciting aspect would be to find an elementary question which is difficult *for a new reason.*) I think that Erdős's distance problems were such questions. Today there is a whole well of deep elementary questions around them. In the next section, we look at some of the other questions in this well, and we try to discuss why they are difficult.

7.5. Open questions

The Szemerédi-Trotter theorem provides a good estimate for

$$\max_{|\mathfrak{L}|=L} |P_r(\mathfrak{L})|.$$

We can pose variations on this question by replacing lines with other types of curves. What happens for circles? Unit circles? Parabolas? Ellipses? These are four interesting open problems of incidence geometry. We do not understand the possible intersection patterns for any of these classes of curves.

Let us try to discuss why these problems are hard, or at least why the methods we have discussed so far do not resolve them. We focus on the example of unit circles, which is probably the example that people have studied the most.

If $\mathfrak{L}$ is a set of L unit circles in the plane, an adaptation of the crossing number proof shows that $|P_r(\mathfrak{L})| \lesssim Lr^{-1} + L^2 r^{-3}$, the same bound as for straight lines. This is the best known upper bound for r-rich points of unit circles, but we don't know examples where this bound is sharp. We can find examples with Lr^{-1} r-rich points by choosing Lr^{-1} points and then choosing r circles through each point. Also, if $r = 2$, it's easy to make examples with $\sim L^2$ 2-rich points. But in the interesting range $3 \leq r \ll L^{1/2}$, there is a big gap between the examples and the upper bound. For straight lines, the grid example gives $\sim L^2 r^{-3}$ r-rich points. But the grid example does not work nearly as well if we replace lines by circles. The

difference is that a circle can contain very few points of a grid. For any $\varepsilon > 0$, for an $S \times S$ grid of points, any circle contains $\leq C(\varepsilon) S^{\varepsilon}$ of the points. For large values of r, the grid construction is still the best known construction, and the number of r-rich points is only slightly more than the trivial example Lr^{-1}.

The most well studied variation of the problem is to estimate the maximum number of incidences between L unit circles and L points. This problem is equivalent to estimating the maximum number of unit distances formed by L points. The arguments coming from the Szemerédi-Trotter theorem show that the number of incidences is $\lesssim L^{4/3}$. The best known examples are based on grids. The number of incidences in these examples is superlinear but $\leq C(\epsilon) L^{1+\epsilon}$ for any $\epsilon > 0$.

Understanding the number of 3-rich points is also interesting. The only upper bound that we know is the trivial bound L^2. Elekes found an arrangement of L unit circles with $\sim L^{3/2}$ 3-rich points [**El1**].

It is widely believed that the maximal number of incidences is close to the examples, at least for the unit distance problem. This means that the upper bounds need to be improved. The upper bound arguments that we know apply to both unit circles and straight lines, and they are sharp for straight lines. To improve these upper bounds, we need to use a property that holds for unit circles but not for straight lines. It is hard to come up with a good candidate for this property. We can get another perspective by comparing unit circles with unit parabolas. A unit parabola is the graph of an equation of the form $y = x^2 + ax + b$, for $a, b \in \mathbb{R}$. The crossing number arguments apply to unit parabolas as well as unit circles and give the same estimates: $|P_r(\mathfrak{L})| \lesssim Lr^{-1} + L^2 r^{-3}$. For unit parabolas, these bounds are sharp up to constant factors, like for straight lines. The reason is that there is a transformation $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ that turns straight lines into unit parabolas. The transformation is $\Phi(x, y) = (x, y + x^2)$. The map Φ is bijective. If l is a (non-vertical) straight line, then $\Phi(l)$ is a unit parabola. If γ is a unit parabola, then $\Phi^{-1}(l)$ is a straight line. Therefore, the incidence geometry of straight lines and of unit parabolas is exactly the same. On the other hand, unit parabolas and unit circles have many features in common. A unit parabola and a unit circle are each degree 2 algebraic curves. They are each strictly convex. They are each determined by two real parameters. Two unit circles intersect in at most two points. Two unit parabolas intersect in at most one point, which seems even better. What special property of unit circles can we use to distinguish them from unit parabolas and from straight lines? (See [**ESS**] for some interesting ideas about this question.)

This discussion suggests another question. Do mathematicians know of any other configurations of lines, besides the grid example, that produce a lot of r-rich points? For small values of r, there are some other examples. For $r = 3$, there is an example based on a degree 3 algebraic curve which was found by Sylvester in the 19th century – see [**GT**] for a discussion of this example. For other small values of r, there are more recent examples – see [**SoSt**] and the references therein. But for r in the range $L^{.01} < r < L^{.49}$, the only known examples that give $\sim L^2 r^{-3}$ r-rich points are minor variations of the grid example. It's striking how few examples we know. It could be that there are more exotic examples we haven't found yet, or it could be that integer grids are essentially the only examples. Understanding this issue is another major open problem in incidence geometry.

We can get a little bit of perspective on this problem by just counting parameters. A line in $\mathbb{R}^2$ is determined by two real parameters. In our questions, there

is no real loss of generality in restricting to lines of the form $y = mx + b$. The space of configurations of L lines $l_j \subset \mathbb{R}^3$ is determined by $2L$ real parameters: $m_j, b_j,$, with $j = 1, ..., L$. The condition that the three lines l_{j_1}, l_{j_2}, l_{j_3} intersect in a common point can be written as one equation in the six variables that define these lines: $m_{j_1}, b_{j_1}, m_{j_2}, b_{j_2}, m_{j_3}, b_{j_3}$. The condition that r lines, $l_{j_1}, ..., l_{j_r}$ all meet in a common point can be written as $r - 2$ equations in the $2r$ variables defining these lines.

Suppose that we have a configuration of L lines in $\mathbb{R}^2$ with N r-rich points. There is some combinatorial information which encodes which lines intersect at each of these r-rich points. Given this combinatorial information, we can write down $(r-2)N$ equations which must be solved by the $2L$ real parameters m_j, b_j representing the set of lines. If $(r-2)N > 2L$, then this system of equations is overdetermined.

Let us look back at the examples in the Szemerédi-Trotter problem with this idea of counting equations and variables. We gave one example with $N = Lr^{-1}$ r-rich points. In this example, $(r-2)N < 2L$, so this example was actually underdetermined. Using generic lines, we gave an example with $N = \binom{L}{2}$ 2-rich points. In this case $(r-2)N = 0 < 2L$, so this example was also underdetermined. As we saw in the discussion on unit circles, these underdetermined examples easily generalize to unit circles. Finally, we gave the grid example, with $N \sim L^2 r^{-3}$ r-rich points. If $3 \leq r \ll L^{1/2}$, then $(r-2)N$ is far bigger than $2L$. So the grid example is heavily overdetermined.

Faced with a heavily overdetermined system of equations, it is a reasonable first guess that it has no solutions at all. If it does have solutions, then we might guess that the problem has some special structure, and that the solutions are all related to the special structure. Just based on this parameter counting, it is plausible to hope that all the near-sharp examples in the Szemerédi-Trotter problem have some special structure. Whether this is really true and exactly what the structure should be is a major open problem.

The cutting of edge of research in this direction is a recent paper of Green and Tao, [**GT**]. Given a set of points $P \subset \mathbb{R}^2$, we let $\mathfrak{L}_{=3}(P)$ be the set of lines containing exactly three points of P. For all sufficiently large N, [**GT**] finds the exact maximum size of $|\mathfrak{L}_{=3}(P)|$ among all sets of N points. This is dual to a problem about lines. If $\mathfrak{L}$ is a set of lines, we let $P_{=3}(\mathfrak{L})$ be the set of points contained in exactly three lines of $\mathfrak{L}$. For all sufficiently large N, [**GT**] finds the exact maximum size of $|P_{=3}(\mathfrak{L})|$ among all sets of N lines. The sharp examples are based on degree 3 algebraic curves. And [**GT**] also shows that if the size of $|\mathfrak{L}_{=3}(P)|$ is very close to the maximum, then P must closely resemble the sharp example.

In Chapter 11 and Chapter 13, we will discuss the connection between combinatorial structure and algebraic structure. For some incidence geometry problems about lines in $\mathbb{R}^3$, we will prove that all the near-sharp examples have a special structure based on low degree polynomials. The special structures for these problems about lines in $\mathbb{R}^3$ turn out to be simpler to understand than the possible special structure for the Szemerédi-Trotter problem. We understand these structure problems a lot better, and they play an important role in the proof of the distinct distance estimate.

7.5.1. Exercises.

EXERCISE 7.10. In this exercise, we describe an example, due to Elekes [**El1**], of a set of N unit circles in the plane with $\sim N^{3/2}$ 3-rich points.

Let $v_1, ..., v_m$ be a set of generic unit vectors in $\mathbb{R}^2$. Let X be the set of all sums $v_i + v_j$ with i and j distinct. If the v_i are chosen generically, then all the sums are distinct, and the number of points in X is $N \sim m^2$. Let Γ be the set of all unit circles with centers in X. Check that every point of the form $v_i + v_j + v_k$ with i, j, k distinct is a 3-rich point of Γ. If the v_i are chosen generically, then these points are all distinct. Conclude that the number of 3-rich points of Γ is $\gtrsim m^3 \sim N^{3/2}$.

EXERCISE 7.11. In [**El2**], Elekes gave an interesting application of the Szemerédi-Trotter theorem to combinatorial number theory. Suppose that $A \subset \mathbb{R}$ is a finite set. We let $A + A$ denote the set of all sums:

$$A + A := \{a_1 + a_2 | a_1, a_2 \in A\}.$$

Similarly, we let $A \cdot A$ denote the set of all products:

$$A \cdot A := \{a_1 \cdot a_2 | a_1, a_2 \in A\}.$$

For a generic set A, $|A + A|$ and $|A \cdot A|$ are both on the order of $|A|^2$. It is interesting to understand how small these sets can be. If A is an arithmetic progression, then $|A + A|$ is roughly $2|A|$. If A is a geometric progression, then $|A \cdot A|$ is roughly $2|A|$. But it seems to be difficult for $|A + A|$ and $|A \cdot A|$ to be small at the same time. Erdős and Szemerédi conjectured that for any $\varepsilon > 0$

$$\max(|A + A|, |A \cdot A|) \geq c(\varepsilon)|A|^{2-\varepsilon}. \tag{7.1}$$

The best known results on this problem are far from the conjecture. In [**El2**], Elekes proved the bound

$$\max(|A + A|, |A \cdot A|) \gtrsim |A|^{5/4}. \tag{7.2}$$

This result is no longer the best known, but it was a big milestone.

Let $\mathfrak{L}$ be the set of lines $y = mx - b$ with $m \in A \cdot A$ and $b \in A + A$. If $1/x \in A$ and $-y \in A$, then check that (x, y) is an r-rich point of $\mathfrak{L}$ for $r = |A|$. Applying Theorem 7.1 to bound $|P_r(\mathfrak{L})|$, prove Inequality 7.2.

Modifying this argument a little bit, prove that $\max(|A - A|, |A/A|) \gtrsim |A|^{5/4}$ also. (Here $A - A$ is the set of differences $a_1 - a_2$, and A/A is the set of quotients a_1/a_2, with $a_1, a_2 \in A$.)

7.6. Crossing numbers and distance problems

In this section, we explore the distinct distance problem and the unit distance problem using the crossing number method. The section mostly consists of exercises where the reader can practice using the crossing number method. Over the course of the exercises, we will prove some interesting results about distance problems, and we will see some of the approaches to these problems that people have tried. We will also try to give a sense of why it is hard to fully solve these problems using the crossing number method. In this way, we will think more about the nature of the difficulty of these problems.

The main tool in this section is the crossing number theorem of [**Le**] and [**ACNS**], Theorem 7.6, which we repeat here.

THEOREM. If G is a graph with E edges and V vertices, and $E \geq 4V$, then the crossing number G obeys the inequality $k(G) \geq (1/64)E^3V^{-2}$.

To begin our exploration, let us sketch a wrong proof of the distinct distance conjecture using the crossing number theorem. As you read the proof, try to figure out where the mistake is.

Suppose that $P \subset \mathbb{R}^2$ is a set of N points with $|d(P)| = t < N$. Consider the set of circles with centers at points of P and radii in $d(P)$. The number of circles in this set is Nt. Each point of P lies in $N-1$ of the circles. Using these circles, we construct a graph G. The vertices of G will be the points of P, and the edges will be the arcs of circles between consecutive points. What is the crossing number of the graph? We know that any two circles intersect in at most two points. So we have the inequality $k(G) \leq 2\binom{Nt}{2} \leq (Nt)^2$. On the other hand, we know that the graph G has N vertices. Each point of P is contained in $N-1$ circles, so each vertex has degree $2(N-1)$. So G has $N(N-1)$ edges. Using the crossing number theorem, we get the following inequalities:

$$N^2t^2 \geq K(G) \geq (1/64)E^3V^{-2} = (1/64)N^3(N-1)^3N^{-2} \geq (1/100)N^4.$$

Solving for t, this implies that $|d(P)| = t \geq (1/10)N$. This inequality is actually wrong if P is a square grid, and the proof must also be wrong. Where is the mistake? Before you read ahead, try to go through the argument carefully, draw a picture, and see if you can find the mistake.

The mistake in the argument is that G is not a graph. In the definition of G above, there may be multiple edges with the same two endpoints. Suppose that $q_1, q_2 \in P$ and that P contains many points on the perpendicular bisector of the segment from q_1 to q_2. Then our set of circles will contain many different circles that go through q_1 and q_2. Potentially, this could create many edges from q_1 to q_2. (In addition to multiple edges, G may contain loops. A loop is an edge whose two endpoints are the same.)

DEFINITION 7.12. We will use the term **multigraph** to refer to a graph that can have multiple edges, but no loops. For a multigraph G, define $Mult(G)$ as the highest number of parallel edges between two points, so $Mult(G) \leq M$ implies that no two points have more than M edges between them.

The crossing number of a multigraph is defined in the same way as for a graph.

EXERCISE 7.12. Prove a crude form of the crossing number theorem for multigraphs with a bound on $Mult(G)$. For instance, the following Proposition has a short proof using the crossing number theorem, Theorem 7.6.

PROPOSITION 7.13. If G is a multigraph with $Mult(G) \leq M$, and $E \geq 4MV$, then $K(G) \geq 1/64E^3V^{-2}M^{-3}$.

EXERCISE 7.13. Modifying the wrong proof above, and using Proposition 7.13, prove the following result about the distinct distance problem.

THEOREM 7.14. If we have N points in the plane, no 100 of which are on a common line, then the number of distinct distances is at least cN, where c is a constant.

Hint: If there are many edges from q_1 to q_2, then there must be many points of P on their perpendicular bisector.

EXERCISE 7.14. Adapt the crossing number proof of Szemerédi-Trotter to the unit distance problem. Using Proposition 7.13, prove the following theorem.

THEOREM 7.15. (Spencer, Szemerédi, Trotter [**SST**]) A set S of N points in the plane determine at most $CN^{4/3}$ unit distances.

Hint: It is crucial to observe that through any two points there at most two unit circles.

This theorem is currently the best known result about the unit distance problem.

EXERCISE 7.15. Examining the proof of the last result and seeing what properties of unit circles really appear, prove the following generalization.

THEOREM 7.16. Suppose that Γ is a set of L connected curves in the plane. (The curves can be either closed like circles or unbounded like lines.) Suppose that any two curves of Γ intersect in at most s points and any two points lie in at most s curves of Γ. Prove that

$$|P_r(\Gamma)| \leq C(s)(L^2 r^{-3} + Lr^{-1}).$$

The last theorem is sharp for either straight lines or unit parabolas. Recall that a unit parabola is given by the equation $y = x^2 + ax + b$.

EXERCISE 7.16. Give an example of a set Γ of N unit parabolas with N r-rich points for $r \gtrsim N^{1/3}$.

Here is another application of this theorem.

EXERCISE 7.17. Suppose that γ_0 is a closed strictly convex curve in the plane. Let Γ be a set of translates of γ_0. Prove that any two points in the plane lie in at most two of the curves of Γ.

Applying Theorem 7.16, conclude that $|P_r(\Gamma)| \lesssim |\Gamma|^2 r^{-3} + |\Gamma| r^{-1}$.

Suppose that γ_0 is a closed strictly convex curve in the plane of diameter d. Prove that the number of integer points on γ_0 is $\lesssim d^{2/3}$. (This result goes back to Jarnik [**Ja**] in the early 20th century. It has a more elementary proof, without topology, but this proof using incidence geometry is pretty.)

Now we return to the crossing numbers of multigraphs and try to prove a sharper estimate for $k(G)$ in terms of the number of edges, the number of vertices, and the multiplicity. We illustrate the issues on an example. Let $K_{5;M}$ be the multigraph with 5 vertices and M edges between each pair of vertices. It has multiplicity M. What is the crossing number of $K_{5;M}$. How does it depend asymptotically on M? We can easily embed $K_{5;M}$ into the plane with M^2 crossings: embed K_5 so that it has one crossing, and draw each edge M times. Can we do better?

Suppose we have a drawing of $K_{5;M}$. Take a random subgraph $G' \subset K_{5;M}$ consisting of one edge between each pair of vertices. In the induced embedding on the subgraph, each crossing occurs with probability $1/M^2$, since it occurs if and only if both edges are in G'. So the number of crossings in the diagram is at least $M^2 \mathbb{E}(k(G')) \geq M^2$. Therefore, $k(K_{5;M}) = M^2$.

EXERCISE 7.18. Generalizing this idea to an arbitrary graph, prove the following result.

THEOREM 7.17. If G is a multigraph with multiplicity at most M, and $E \geq 100MV$, then $K(G) \geq cE^3V^{-2}M^{-1}$ for some c.

Hint: Start with the case that between any two vertices of G the number of edges is either 0 or lies in $[M/2, M]$.

Next, following [**Sz**], we apply the crossing number therorem (for multigraphs) to the distinct distance problem. Using Theorem 7.17, we will prove the following estimate, which was the best known estimate for the distinct distance problem in the late 90's.

THEOREM 7.18. (Székely, [**Sz**]) If we have N distinct points in the plane, then they determine $\geq cN^{4/5}$ distinct distances. In fact, there is one point p in the set so that the distance from p takes $\geq cN^{4/5}$ distinct values.

We describe the proof, writing out some of the ideas and leaving other steps as exercises.

Suppose that for each point p in our set $\mathcal{S}$, the set of distances $\{dist(p,q)\}_{q\in\mathcal{S}}$ takes on $\leq t$ different values. We assume $t \leq cN^{4/5}$ and we will get a contradiction.

We let Γ_0 be the set of circles $S(p,r)$ with centers $p \in \mathcal{S}$, and radii $r = dist(p,q)$ for some $q \in \mathcal{S}$. The total number of circles in Γ_0 is at most Nt. We let $\Gamma \subset \Gamma_0$ be the set of circles in Γ_0 that contain at least two distinct points of $\mathcal{S}$. Using Γ and $\mathcal{S}$, we define a multigraph G as follows. The vertices of G are the points of $\mathcal{S}$ and the edges are arcs between consecutive points on one of the circles of Γ. By leaving out circles that contain only one point, we arrange that G has no loops.

The multigraph G has N vertices and approximately N^2 edges. We estimate the number of edges in the following exercise.

EXERCISE 7.19. If N is large enough, we can assume that $t \leq (1/100)N$. With this assumption, prove that the number of edges of G is at least $(1/2)N^2$.

The multigraph G was constructed with a drawing in the plane, and every crossing in this drawing corresponds to the intersection of two of the circles of Γ. A pair of circles intersects in at most two points, and $|\Gamma| \leq Nt$, and so G has crossing number $\leq 2(Nt)^2$.

The multigraph G may have very high multiplicity. Our strategy will be to estimate how many high-multiplicity edges G can have, and trim edges from G to reduce the multiplicity.

LEMMA 7.19. The number of edges of G with multiplicity $\geq M$ is at most $C[N^2M^{-2}t + N\log Nt]$.

This is one of the harder steps. The reader may certainly find the proof on their own, but we also include a proof here.

PROOF. Consider edges from a vertex p_1 to a vertex p_2. Each edge is the arc of a circle, and the center of the circle must lie on the perpendicular bisector of p_1 and p_2. If there are many edges from p_1 to p_2, then there must be many points of our set along the perpendicular bisector.

We define a map from edges of our multigraph to lines, sending an edge to the corresponding perpendicular bisector. A line containing A points of $\mathcal{S}$ contributes $\leq 2At$ edges of the multigraph, each with multiplicity $\leq A$.

Let $\mathfrak{L}_j$ denote the set of lines in the plane which contain $\sim 2^j$ points of $\mathcal{S}$. (More precisely, the number of points is greater than 2^{j-1} and at least 2^j.) The number of edges with multiplicity at least M is bounded by

$$\sum_{2^j \geq M} |\mathfrak{L}_j| 2 \cdot 2^j t.$$

The size of $\mathfrak{L}_j$ is bounded by the Szemerédi-Trotter theorem (see Version 3 above). Plugging in, we get:

$$\leq \sum_{2^j \geq M} C(N^2 2^{-3j} + N 2^{-j}) 2^j t.$$

The $N^2 2^{-3j}$ term decays exponentially in j, and the total is $\leq CN^2 M^{-2} t$. The second term is independent of j, and we need to sum over $\sim \log N$ values of j, so the total is $\leq CN \log N t$. □

For any M, we define $G_{\leq M} \subset G$ to be the multigraph given by deleting all edges of G with multiplicity $\geq M$. (Recall that the multiplicity of an edge in a multigraph is the number of edges with the same endpoints as the given edge.) For any M, we know that

$$k(G_{\leq M}) \leq k(G) \leq 2(Nt)^2.$$

We want to choose M as small as possible, but still guaranteeing that $G_{\leq M}$ has $\geq (1/3)N^2$ edges. Using Lemma 7.19 we can estimate how small we can make M.

Exercise 7.20. Applying Theorem 7.17 to $G_{\leq M}$, prove that $t \geq cN^{4/5}$ for a constant $c > 0$.

We end this section with a few comments about crossing numbers and distinct distances. Building on the crossing number approach introduced in [**Sz**], Solymosi-Toth [**SoTo**] and then Katz-Tardos [**KatTar**] improved the estimates in the distinct distance problem. The paper [**KatTar**] proved that for any N points in the plane, one of the points determines $\geq cN^{.864}$ distances with the other points. This approach gave the best estimate in the distinct distance problem before the polynomial method approach.

We saw very early on that if a set of N points has at most 100 points on any line, then it has at least cN distinct distances. Therefore, any example with far less than N distinct distances must have a lot of points on a line. If there was only one rich line, then we could deal with it separately, so we are really worried about examples with many rich lines. This sounds like a very special structure for the set of points. At first sight, it seemed to me that this structure should give us a lot of leverage. The argument by [**Sz**] does exploit this structure to some extent. However, this situation recalls one of the basic issues that we discussed in the open questions section, Section 7.5. We know a few examples of sets of points with many rich lines, and these examples have a lot of structure. On the other hand, we can prove very little about the structure of a set of points with many rich lines.

Using the polynomial method we will prove during the book that the number of distinct distances given by N points is $\geq cN(\log N)^{-1}$. However, this approach does

not bound the number of distances from a single point. It looks completely plausible that for any N points in the plane, one of the points determines $\geq cN(\log N)^{-1}$ (or even $\geq cN(\log N)^{-1/2}$)) distances with the other points. This would be a better theorem if it's true.

https://doi.org/10.1090//ulect/064/08

CHAPTER 8

Incidence geometry in three dimensions

In the last chapter, we discussed incidence geometry in the plane. Now we turn to higher dimensions. The polynomial method has led to some significant progress in incidence geometry in higher dimensions, and this is the main topic for the rest of the book. In particular, we will study in depth the incidence geometry of lines in $\mathbb{R}^3$. The joints problem concerns the incidence geometry of lines in $\mathbb{R}^3$, so we have already seen how the polynomial method plays a role.

In the first section, we discuss the incidence geometry of lines in $\mathbb{R}^3$ and we formulate the main results about them that we will prove over the course of the book, using the polynomial method. In the next section, we discuss what is known about even higher dimensions, giving some references to the literature. This area is only beginning to be explored.

After that, we discuss a couple of other topics that play an important role in higher dimensional incidence geometry. The first tool is the Zarankiewicz problem – a fundamental combinatorial problem that comes up in many places in incidence geometry. After that, we introduce reguli. These are degree 2 algebraic surfaces that play an important role in studying lines in three-dimensional space. These are some of the main tools that were used to study incidence geometry of lines in $\mathbb{R}^3$ before the polynomial method, and we give a sample argument showing how they can be applied.

8.1. Main results about lines in $\mathbb{R}^3$

In this section, we consider the incidence geometry of lines in $\mathbb{R}^3$. We might start with the first question we considered in $\mathbb{R}^2$: given a set $\mathfrak{L}$ of L lines in $\mathbb{R}^3$, what is the maximum possible number of r-rich points? It turns out the answer is exactly the same as in the plane.

PROPOSITION 8.1. Suppose $n \geq 2$. Then for any L, r,

$$\max_{\mathfrak{L} \text{ a set of } L \text{ lines in } \mathbb{R}^n} |P_r(\mathfrak{L})| = \max_{\mathfrak{L} \text{ a set of } L \text{ lines in } \mathbb{R}^2} |P_r(\mathfrak{L})| \lesssim L^2 r^{-3} + L r^{-1}.$$

PROOF. Consider a set of lines $\mathfrak{L}$ in $\mathbb{R}^n$. Let $\pi : \mathbb{R}^n \to \mathbb{R}^2$ be a projection. For a generic choice of π, the images of the lines of $\mathfrak{L}$ will be distinct lines in $\mathbb{R}^2$. So $\pi(\mathfrak{L})$ will be a set of L (distinct) lines in $\mathbb{R}^2$. For a generic choice of π, the images of $P_r(\mathfrak{L})$ will be distinct points in $\mathbb{R}^2$, and we always have $P_r(\pi(\mathfrak{L})) \subset \pi(P_r(\mathfrak{L}))$. Therefore, $|P_r(\mathfrak{L})| \leq P_r(\pi(\mathfrak{L}))|$.

On the other hand, $\mathbb{R}^2 \subset \mathbb{R}^n$, so a set of lines in $\mathbb{R}^2$ can be seen as a set of lines in $\mathbb{R}^n$. This proves that

$$\max_{\mathfrak{L} \text{ a set of } L \text{ lines in } \mathbb{R}^n} |P_r(\mathfrak{L})| = \max_{\mathfrak{L} \text{ a set of } L \text{ lines in } \mathbb{R}^2} |P_r(\mathfrak{L})|.$$

Now by the Szemerédi-Trotter theorem, we see that in any dimension n,

$$|P_r(\mathfrak{L})| \lesssim L^2 r^{-3} + L r^{-1}.$$

□

This result answers the question about the maximal number of r-rich points for L lines in $\mathbb{R}^3$. It turns out that this question is not "really 3-dimensional". The worst examples are when all lines lie in a plane, and the general case quickly reduces to the planar case. It takes some thought to formulate interesting questions about lines in $\mathbb{R}^3$ – questions that are really 3-dimensional and don't reduce to the planar case.

The joints problem is one such question. We now have a little more context to appreciate the joints problem. The joints problem is a question about the incidence geometry of lines in $\mathbb{R}^3$ that doesn't reduce to a 2-dimensional question. It is probably the simplest really 3-dimensional problem about the incidence geometry of lines in $\mathbb{R}^3$.

Here is another approach to formulating a really 3-dimensional question. We consider a set of lines in $\mathbb{R}^3$ with an extra condition that not too many lines lie in any plane. Under this extra condition, we can ask whether there is a better estimate for $|P_r(\mathfrak{L})|$.

We considered this type of problem in Chapter 3. We saw that a set of L lines in $\mathbb{R}^3$ lying in a degree 2 algebraic surface can have $\sim L^2$ 2-rich points, even though at most two of the lines lie in a plane. This example suggests that maybe we should consider low degree algebraic surfaces as well as planes, leading to the following question.

Question 8.2. Suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ which contains at most B lines in any algebraic surface of degree $\leq D$. What is the maximal possible size of $|P_r(\mathfrak{L})|$?

This type of question came out of work by Elekes-Sharir [**ElSh**]. They gave a new approach to the distinct distance problem, which led to this problem. We will explain this approach in Chapter 9. The connection to distinct distances in their work is an important motivation to look at this question. But I think this question is also a natural question in its own right. I think it does a good job of getting at really 3-dimensional phenomena in the incidence geometry of lines in $\mathbb{R}^3$. The next six chapters of the book are concerned with this question and its applications.

We discuss some examples, and then state our main results about the question. We start by considering 2-rich points. In Section 3.5, we described a configuration of L lines in a certain degree 2 algebraic surface with $\sim L^2$ 2-rich points. The degree 2 surface was defined by the equation $z = xy$. It is an example of a regulus, and we will study reguli systematically in Section 8.4 below. In any regulus, for any $L \geq 2$, we can find L lines with $\sim L^2$ 2-rich points.

If we let $S_1, ..., S_{L/B}$ be planes or reguli, and if $\mathfrak{L}$ is a set of lines containing B lines from each S_j, then the number of 2-rich points of $\mathfrak{L}$ can be $\sim (L/B)B^2 \sim LB$. Our first theorem shows that this example is sharp up to a constant factor if $B \geq L^{1/2}$.

Theorem 8.3. If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with at most B lines in any plane or degree 2 surface, then $|P_2(\mathfrak{L})| \lesssim LB + L^{3/2}$.

If $B \geq L^{1/2}$, then our upper bound is $|P_2(\mathfrak{L})| \lesssim LB$, which matches the example above. If $B \leq L^{1/2}$, then the our upper bound is still $|P_2(\mathfrak{L})| \lesssim L^{3/2}$. Even for small B, the only examples that I currently know have $|P_2(\mathfrak{L})| \lesssim BL$.

Now we consider r-rich points for $r \geq 3$. Suppose we take L/B planes, and we let $\mathfrak{L}$ be a set containing B lines in each of the planes. The grid example from Section 7.1 is a configuration of B lines in a plane with $\sim B^2 r^{-3}$ r-rich points. If we use a grid example in each plane, then the total number of r-rich points will be $(L/B)B^2 r^{-3} = BLr^{-3}$. The value $B = L^{1/2}$ will be an important example in the book, and in this case, we note that the number of r-rich points is $L^{3/2}r^{-3}$.

Here is another example that is more 3-dimensional. Let G_0 denote the integer lattice $\{(a, b, 0)\}$ with $1 \leq a, b \leq L^{1/4}$. Let G_1 denote the integer lattice $\{(a, b, 1)\}$ with $1 \leq a, b \leq L^{1/4}$. Let $\mathfrak{L}$ denote all the lines from a point of G_0 to a point of G_1. Since G_0, G_1 each contain $L^{1/2}$ points, $\mathfrak{L}$ is a set of L lines. The horizontal planes $z = 0$ and $z = 1$ do not contain any lines of $\mathfrak{L}$. Any other plane contains at most $L^{1/4}$ points of each G_i, and so at most $L^{1/2}$ lines of $\mathfrak{L}$. In this example, $|P_r(\mathfrak{L})| \sim L^{3/2}r^{-2}$ for all r in the range $2 \leq r \leq L^{1/2}/400$. We put this computation in the exercises with some guidance. Notice that for $B = L^{1/2}$, and for large r, this example has more r-rich points than the previous example. Our second main theorem says that when $B = L^{1/2}$, this example is optimal.

THEOREM 8.4. If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with at most $L^{1/2}$ lines in any plane, and if $3 \leq r \leq 2L^{1/2}$, then $|P_r(\mathfrak{L})| \lesssim L^{3/2}r^{-2}$.

(If $r \geq 2L^{1/2}$ then $|P_r(\mathfrak{L})| \leq 2Lr^{-1}$. We proved this for lines in the plane in Lemma 7.3, and the same argument applies in any dimension.)

The proof of Theorems 8.3 and 8.4 will take some work. In the rest of this chapter, we give some more background about incidence geometry in dimension at least three. In the next chapter, we explain the connection between these estimates about lines in $\mathbb{R}^3$ and the distinct distance problem. In Chapters 10 to 13, we prove Theorems 8.3 and 8.4. We will prove a slightly weaker estimate at the end of Chapter 10. We will finish the proof of Theorem 8.4 at the end of Chapter 12 and we will finish the proof of Theorem 8.3 at the end of Chapter 13.

EXERCISE 8.1. In this exercise, we estimate the number of r-rich points for the set of lines described a few paragraphs ago. Namely, let G_0 be the set of points $(a, b, 0)$ with a, b integers in the range $1 \leq a, b \leq L^{1/4}$, and G_1 be the set of points $(a, b, 1)$ with a, b integers in the same range. Let $\mathfrak{L}$ be the set of all lines that contain one point of G_0 and one point of G_1. The number of lines in $\mathfrak{L}$ is L. The goal of the exercise is to show that for all r in the range $2 \leq r \leq (1/400)L^{1/2}$, we have the estimate

$$|P_r(\mathfrak{L})| \gtrsim L^{3/2}r^{-2}.$$

Here is an approach to studying the r-rich points of $\mathfrak{L}$. For any point $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, with $x_3 \neq 0, 1$, we define a map $\rho_x : \mathbb{R}^2 \to \mathbb{R}^2$ as follows. If $v \in \mathbb{R}^2$, then we define $\rho_x(v)$ so that $(v, 0)$, x, and $(\rho_x(v), 1)$ are collinear. Now observe that x is r-rich if and only if

$$|\rho_x(G_0) \cap G_1| \geq r.$$

Using a similar triangles argument, show that

$$\rho_{(0,0,x_3)})(v) = -\frac{1 - x_3}{x_3} v.$$

More generally, show that if $x = (x_1, x_2, x_3)$,

$$\rho_x(v_1, v_2) = -\frac{1 - x_3}{x_3}(v - (x_1, x_2)) + (x_1, x_2) = -\frac{1 - x_3}{x_3} v + \frac{1}{x_3}(x_1, x_2).$$

If p/q is a rational number (in lowest terms) with $\max(|p|, |q|) \sim H$, show that there are $\gtrsim H^2 L^{1/2}$ choices of (x_1, x_2) so that $(x_1, x_2, p/q)$ lies in $\geq (1/10) H^{-2} L^{1/2}$ lines of $\mathfrak{L}$.

Adding up the contributions from different rational heights, show that $|P_r(\mathfrak{L})| \gtrsim L^{3/2} r^{-2}$ for all r in the range $2 \leq r \leq (1/400) L^{1/2}$.

8.2. Higher dimensions

So far we have discussed in some depth the incidence geometry of lines in $\mathbb{R}^2$ and of lines in $\mathbb{R}^3$. The incidence geometry of lines in $\mathbb{R}^3$ will be the main subject of the rest of the book. It would be very interesting to have a systematic theory of the incidence geometry of k-planes in $\mathbb{R}^n$ for every k, n. So far we know only a little in this direction. It's not even clear to me what questions a systematic theory should aim to answer.

Let us first consider 2-planes in $\mathbb{R}^3$. We might first ask, given a set Π of N 2-planes in $\mathbb{R}^3$ and a set $\mathcal{S}$ of S points in $\mathbb{R}^3$, how many incidences can they form? The answer to this question is NS – in other words, every point can lie in every plane. To achieve this, pick a line $l \subset \mathbb{R}^3$, and then choose N planes containing l and S points in l. To get interesting bounds on the number of incidences, it is necessary to put further restrictions on the planes and/or the points.

Mathematicians have considered many different possible restrictions. Edelsbrunner, Guibas, and Sharir [**EGS**] consider the restriction that any line contains at most two points of $\mathcal{S}$. Chazelle [**Ch**] considers the restriction that there is no line which contains at least r points of $\mathcal{S}$ and lies in at least r planes of Π. Agarwal and Aronov [**AA**] consider the restriction that each plane in Π contains three non-collinear points of $\mathcal{S}$. Elekes and Toth [**ElTo**] consider a stronger version of this restriction: for each plane $\pi \in \Pi$, at most $\alpha |\mathcal{S} \cap \pi|$ points of $\mathcal{S} \cap \pi$ can lie on any line, where $0 < \alpha < 1$ is given. Under these different restrictions, there are various interesting bounds on the number of incidences. Most of these bounds are not known to be sharp, although there are a few sharp inequalities known. These different setups provide complementary estimates – there doesn't seem to be one central conjecture/question about incidences of points and 2-planes.

To give a flavor for the area, we state one theorem about incidences of points and planes due to [**EGS**]. (See also Appendix A of [**ApSh**] for the sharpest version and some interesting discussion.)

THEOREM 8.5. Suppose that Γ is a set of N 2-planes in $\mathbb{R}^3$, no three of which are collinear. Suppose that $\mathcal{S}$ is a set of S points in $\mathbb{R}^3$. Then

$$|I(\Gamma, \mathcal{S})| \lesssim S^{4/5} N^{3/5} + S + N.$$

This inequality is sharp, as explained in Appendix A of [**ApSh**]. We will study incidences between points and planes in the exercises in Chapter 10. In particular, in Exercise 10.6, we will prove a slightly weaker version of Theorem 8.5.

Some of the work on 2-planes in $\mathbb{R}^3$ generalizes to hyperplanes in $\mathbb{R}^n$ for any dimension n – see [**Ch**] and [**ApSh**]. Point-hyperplane incidences are already quite complex. But even more generally, we would like to understand incidences between

points and k-planes in $\mathbb{R}^n$ for any k, n. Dealing with other values of k creates new problems. In Chapter 10, we will discuss the cutting method, an important method for studying incidences between points and hyperplanes, and we will discuss why it is hard to adapt to k-planes for $k < n-1$.

For k-planes in $\mathbb{R}^n$, with arbitrary k, n, I know of one interesting setup where we have a good estimate for the number of incidences. This setup was proposed by Toth in the early 2000's. Suppose that Π is a set of k-planes in $\mathbb{R}^n$ with the extra condition that any two k-planes of Π intersect in at most one point. Let $\mathcal{S}$ be a set of points in $\mathbb{R}^n$. Given $|\Pi|$ and $|\mathcal{S}|$, what is the maximum possible number of incidences between Π and $\mathcal{S}$?

The condition that two k-planes intersect in at most one point requires $n \geq 2k$. The case $n = 2k$ is the main case. By a random projection argument, as in Proposition 8.1, the case $n = 2k$ is equivalent to any other ambient dimension $n \geq 2k$.

We can get interesting examples by taking products of (interesting) configurations of lines in the plane. For $j = 1, ..., k$, suppose that $\mathfrak{L}_j$ is a set of lines in $\mathbb{R}^2$ and $\mathcal{S}_j$ is a set of points in $\mathbb{R}^2$. Define $\mathcal{S}$ to be the product $\mathcal{S}_1 \times ... \times \mathcal{S}_k \subset \mathbb{R}^{2k}$, so that $|\mathcal{S}| = \prod_{j=1}^k |\mathcal{S}_j|$. Similarly, we define Π to be the product $\mathfrak{L}_1 \times ... \times \mathfrak{L}_k$. An element of $\mathfrak{L}$ is a k-plane in $\mathbb{R}^{2k}$ defined as a product $l_1 \times ... \times l_k \subset \mathbb{R}^{2k}$, where $l_j \in \mathfrak{L}_j$. So Π is a set of $\prod_{j=1}^k |\mathfrak{L}_j|$ k-planes in $\mathbb{R}^{2k}$. The number of incidences of Π and $\mathcal{S}$ is $\prod_{j=1}^k |I(\mathfrak{L}_j, \mathcal{S}_j)|$. For instance, suppose that for each j, $|\mathfrak{L}_j| = |\mathcal{S}_j| = N$ and $I(\mathfrak{L}_j, \mathcal{S}_j) \sim N^{4/3}$. Then $|\Pi| = |\mathcal{S}| = N^k$ and $|I(\Pi, \mathcal{S})| \sim (N^k)^{4/3}$. Toth conjectured that this example is sharp up to a constant factor.

Toth raised this question in connection with his work on the complex version of the Szemerédi-Trotter theorem. Suppose that $\mathfrak{L}$ a set of complex lines in $\mathbb{C}^2$, and $\mathcal{S}$ a set of points in $\mathbb{C}^2$. In terms of $|\mathfrak{L}|$ and $|\mathcal{S}|$, what is the maximum possible number of incidences between $\mathfrak{L}$ and $\mathcal{S}$? For real lines, this question is answered by the Szemerédi-Trotter theorem. But the topological methods used for real lines, such as the crossing number method, do not easily adapt to the complex setting. Toth was able to adapt the original proof from [**SzTr**] to the complex setting, proving the following result.

THEOREM 8.6. (Toth, [**To**]) Suppose that $\mathfrak{L}$ a set of complex lines in $\mathbb{C}^2$, and $\mathcal{S}$ a set of points in $\mathbb{C}^2$. Then

$$|I(\mathfrak{L}, \mathcal{S})| \lesssim \max(|\mathfrak{L}|^{2/3}|\mathcal{S}|^{2/3}, |\mathfrak{L}|, |\mathcal{S}|).$$

Along the way, he observed that complex lines can be thought of as real 2-planes in $\mathbb{R}^4$, where every pair of 2-planes intersects in at most one point. This suggested to him the question above, and he made the following conjecture.

CONJECTURE 8.7. (Toth) Suppose that Π is a set of k-planes in $\mathbb{R}^n$ with the extra condition that any two k-planes of Π intersect in at most one point. Let $\mathcal{S}$ be a set of points in $\mathbb{R}^n$. Then

$$|I(\Pi, \mathcal{S})| \lesssim_k \max(|\Pi|^{2/3}|\mathcal{S}|^{2/3}, |\Pi|, |\mathcal{S}|).$$

Because of the product examples, this upper bound would be sharp up to a constant factor.

Using the polynomial method, Solymosi and Tao were able to prove an estimate which nearly gives Toth's conjecture.

THEOREM 8.8. (Solymosi-Tao, [**SoTa**]) For any $\varepsilon > 0$, and any $k \geq 1$, there is a constant $C(k, \varepsilon)$ so that the following holds. Suppose that Π is a set of k-planes in $\mathbb{R}^n$ with the extra condition that any two k-planes of Π intersect in at most one point. Let $\mathcal{S}$ be a set of points in $\mathbb{R}^n$. Then

$$|I(\Pi, \mathcal{S})| \leq C(k, \varepsilon)|\Pi|^{\varepsilon} \max(|\Pi|^{2/3}|\mathcal{S}|^{2/3}, |\Pi|, |\mathcal{S}|).$$

We don't prove this theorem in this book, but it is related to many of the ideas in the book. In particular, polynomial partitioning, the subject of Chapter 10, plays an important role in the argument. After reading Chapter 10, the reader will be ready to read this interesting paper.

8.3. The Zarankiewicz problem

There is a fundamental combinatorial problem called the Zarankiewicz problem that comes up in several places in incidence geometry.

To motivate the problem, we start by considering incidences between points and lines in the plane. Suppose that $\mathcal{S}$ is a set of points in the plane and $\mathfrak{L}$ is a set of lines in the plane. The incidence matrix of $(\mathcal{S}, \mathfrak{L})$ encodes which points lie on which line. It has one row for each point of $\mathcal{S}$ and one column for each line of $\mathfrak{L}$. Given a point $x \in \mathcal{S}$ and a line $l \in \mathfrak{L}$, the corresponding entry of the matrix is 1 if $x \in l$ and 0 otherwise. The number of 1's in the incidence matrix is the number of incidences between $\mathcal{S}$ and $\mathfrak{L}$.

A fundamental question in incidence geometry is to understand which 0-1 matrices occur as the incidence matrix of some pair $(\mathcal{S}, \mathfrak{L})$. Because two lines intersect in at most one point, the incidence matrix $(\mathcal{S}, \mathfrak{L})$ contains no 2×2 minor of all 1's. This basic fact leads to some estimates on the number of incidences $I(\mathcal{S}, \mathfrak{L})$. In this section, we will study this type of estimate systematically.

In the early 1950's, Zarankiewicz posed the following more general problem. Suppose that A is an $M \times N$ matrix with entries 0 or 1, and suppose that A has no $V \times W$ minor of all 1's. What is the maximum possible number of 1's that A can have? This is a fundamental problem of combinatorics, and various special cases occur in incidence geometry in different places. The main known result is the following theorem.

THEOREM 8.9. (Kővári-Sós-Turán, 1954) Suppose that A is an $M \times N$ matrix whose entries are 0 or 1. Suppose that A has no $V \times W$ minor of all 1's, for some integers $V \leq W$. Then the number of 1's in A is at most $C(V)[W^{1/V} M N^{\frac{V-1}{V}} + N]$.

Remark. We can write the upper bound above in the less precise form $C(V, W)[MN^{\frac{V-1}{V}} + N]$. Sometimes, we have V fixed, while W tends to infinity with M, N. In this case, the more explicit dependence on W is useful.

PROOF. Let $C_1, ..., C_N$ denote the columns of A. We can think of each column as a subset of the numbers $[1, ..., M]$. We let $\binom{M}{V}$ denote all of the sets of V distinct elements of the numbers $1, ..., M$. We let $\binom{C_j}{V}$ denote all of the sets of V distinct elements of C_j. Clearly $\binom{C_j}{V} \subset \binom{M}{V}$. We let $|C_j|$ be the number of elements in C_j, so that the number of elements in $\binom{C_j}{V}$ is $\binom{|C_j|}{V}$.

The condition that A has no $V \times W$ minor of all 1's implies that each element of $\binom{M}{V}$ occurs in $< W$ of the sets $\binom{C_j}{V}$. So we get the following inequality:

$$\sum_{j=1}^{N} \binom{|C_j|}{V} < W \binom{M}{V}.$$

The expression $\binom{|C_j|}{V}$ is somewhat complicated, but it's approximately equal to $|C_j|^V$. We write $A \lesssim B$ for $A \leq C(V)B$. Then $|C_j|^V \lesssim \binom{|C_j|}{V} + 1$. We need the $+1$ term in case $1 \leq |C_j| \leq V - 1$. Plugging this in, we get

$$\sum_{j=1}^{N} |C_j|^V \lesssim \sum_{j=1}^{N} (\binom{|C_j|}{V} + 1) \leq WM^v + N.$$

The total number of 1's in A is $\sum_j |C_j|$. Now by Holder's inequality,

$$\sum_{j=1}^{N} |C_j| \leq (\sum |C_j|^V)^{1/V} N^{\frac{V-1}{V}} \lesssim (WM^V + N)^{1/V} N^{\frac{V-1}{V}} \lesssim W^{1/V} M N^{\frac{V-1}{V}} + N.$$

□

Returning to the incidence matrix of $\mathcal{S}$ and $\mathfrak{L}$, we get the following corollary.

COROLLARY 8.10. If $\mathcal{S}$ is a set of S points and $\mathfrak{L}$ is a set of L lines, then

- $I(\mathcal{S}, \mathfrak{L}) \lesssim SL^{1/2} + L$.
- $I(\mathcal{S}, \mathfrak{L}) \lesssim LS^{1/2} + S$.

PROOF. The incidence matrix of $\mathcal{S}$ and $\mathfrak{L}$ is an $S \times L$ matrix with no 2×2 minor. By Theorem 8.9, the number of 1's is $\lesssim SL^{1/2} + L$. The transpose matrix also has no 2×2 minor, and so the number of 1's is $\lesssim LS^{1/2} + S$ as well. □

Here are a couple other applications of Theorem 8.9 in incidence geometry. None of these bounds are sharp, but they still play a useful role in incidence geometry.

COROLLARY 8.11. If $\mathcal{S}$ is a set of S points in $\mathbb{R}^2$ and Γ is a set of L unit circles, then the number of incidences obeys the following bounds:

- $I(\mathcal{S}, \Gamma) \lesssim SL^{1/2} + L$.
- $I(\mathcal{S}, \Gamma) \lesssim LS^{2/3} + S$.

PROOF. The incidence matrix of $\mathcal{S}$ and Γ is an $S \times L$ matrix. Three unit circles can intersect in at most one point, and so the matrix has no 2×3 minor of all 1's. By Theorem 8.9, the number of incidences is $\lesssim SL^{1/2} + L$ as above. If we take the transpose, we see that the number of incidences is also $\lesssim LS^{2/3} + S$. □

In the unit distance problem, we have N points which are the centers of N unit circles. The number of unit distances is the number of incidences between the points and the unit circles. The argument above shows that the number of unit distances is $\lesssim N^{3/2}$. This argument appears in the first paper on the unit distance problem, [**Er1**].

Now consider S points and L unit spheres in $\mathbb{R}^3$.

COROLLARY 8.12. If $\mathcal{S}$ is a set of S points in $\mathbb{R}^3$ and Γ is a set of L unit spheres, then the number of incidences obeys the following bounds:

- $I(\mathcal{S}, \Gamma) \lesssim SL^{2/3} + L$.
- $I(\mathcal{S}, \Gamma) \lesssim LS^{2/3} + S$.

PROOF. The incidence matrix of $\mathcal{S}$ and Γ is an $S \times L$ matrix. Three unit spheres intersect in at most 2 points, and so the incidence matrix has no 3×3 minor of all 1's. By Theorem 8.9, the number of incidences is at most $SL^{2/3} + S$. Taking the transpose matrix, the number of incidences is also at most $LS^{2/3} + L$. □

In particular, this corollary shows that the number of unit distances determined by N points in $\mathbb{R}^3$ is $\lesssim N^{5/3}$.

It's a very interesting question how sharp the Kővári-Sós-Turán theorem is. There are a few cases where the theorem is known to be sharp, but in general this is a deep open problem. To keep our discussion simpler, we focus on square $N \times N$ matrices and square $V \times V$ minors.

Example 1. Consider an $N \times N$ 0-1 matrix with no 2×2 submatrix. The KST theorem says that the matrix has $\lesssim N^{3/2}$ 1's. This estimate is sharp. The example was discovered by Reiman in [**Re**]. We have essentially already seen the example: it is the incidence matrix of lines over a finite field. We pick $N = q^2$ lines in the plane $\mathbb{F}_q^2$. We let the rows of our matrix correspond to the points of $\mathbb{F}_q^2$ and the columns correspond to the q^2 chosen lines. We put a 1 in the matrix if the point corresponding to the row lies in the line corresponding to the column. Since two lines intersect in at most 1 point, there are no 2×2 submatrices. Since each line contains q points, our matrix has $q^3 = N^{3/2}$ 1's. Working with the projective plane over $\mathbb{F}_q$ is even better: it gives an example with exactly the maximum possible number of 1's.

Example 2. Next consider an $N \times N$ 0-1 matrix with no 3×3 submatrix. The Kővári-Sós-Turán theorem says that the matrix has $\lesssim N^{5/3}$ 1's. In the early 60's, Brown gave an example with $\gtrsim N^{5/3}$ 1's. Brown's construction uses 'spheres' of a fixed radius over a finite field. Suppose that $x \in \mathbb{F}_q^3$ and $r \in \mathbb{F}_q$. We define the 'sphere' $S(x, r)$ as follows:

$$S(x, r) := \{y \in \mathbb{F}_q^3 | \sum_i (x_i - y_i)^2 = r\}.$$

For a fixed r, there are $q^3 = N$ spheres $S(x, r) \subset \mathbb{F}_q^3$, and there are $N = q^3$ points $y \in \mathbb{F}_q^3$. Brown's matrix is the incidence matrix of these spheres and points for well-chosen values of q and r. For many q, r, the sphere $S(0, r)$ has $\sim q^2 = N^{2/3}$ points. Since each point $y \in \mathbb{F}_q^3$ belongs to exactly one sphere $S(0, r)$, the average size of $|S(0, r)| = q^2$, and it turns out there are many r that give close to the average value. For any point x, $S(x, r)$ is just a translate of $S(0, r)$ and so $|S(x, r)| = |S(0, r)|$. Therefore, it is not hard to choose q, r so that the incidence matrix has $\sim N^{5/3}$ 1's.

Now in Euclidean space $\mathbb{R}^3$, it is easy to check that any three spheres of the same radius intersect in at most two points. (Three spheres with different radii may intersect in a circle.) It is not obvious whether this result extends to spheres $S(x, r) \subset \mathbb{F}_q^3$. Brown checked that for some values of q and r this is indeed true. We discuss this issue more in the exercises.

Brown's example is clever and special, and no one knows how to generalize it to 4×4 minors. Consider an $N \times N$ 0-1 matrix with no 4×4 submatrix. The Kővári-Sós-Turán theorem says that the matrix has $\lesssim N^{7/4}$ 1's. The best known examples have only $\sim N^{5/3}$ 1's - and these are Brown's examples which have no 3×3 minor of all 1's! It's a longstanding open problem in combinatorics where the truth lies between Brown's example and the Kővári-Sós-Turán upper bound.

Example 3. Finally, consider an $N \times N$ 0-1 matrix with no $V \times V$ submatrix. The Kővári-Sós-Turán theorem says that the number of 1's is $\leq C(V)N^{2-(1/V)}$. For $V \geq 6$, the best known examples come from a random construction.

PROPOSITION 8.13. Fix V. For each N, there is an $N \times N$ 0-1 matrix with no $V \times V$ minor of all 1's and with $\geq c(V)N^{2-\frac{2}{V+1}}$ 1's.

PROOF. Let p be a probability to be chosen later. We assign the entries of an $N \times N$ matrix independently, giving each entry a 1 with probability p and a 0 with probability $1-p$. The expected number of 1's in the matrix is $N^2 p$.

The expected number of $V \times V$ minors of all 1's is $\binom{N}{V}^2 p^{V^2}$. We call a $V \times V$ minor of all 1's a bad minor. We choose a random matrix M and then delete a 1 from each of its bad minors. The resulting matrix has no bad minors, and the number of 1's left is at least the number of 1's in M minus the number of bad minors in M. By an averaging argument, we can choose a matrix M so that the number of 1's minus the number of bad $V \times V$ minors is at least the expected value

$$N^2 p - \binom{N}{V}^2 p^{V^2} \geq N^2 p - N^{2V} p^{V^2}.$$

Finally, we choose $p \in [0,1]$ to maximize the right hand side. Taking $p = (1/2)N^{-\frac{2}{V+1}}$ gives the result. □

(If $V = 4$, then Brown's construction beats the random construction. For $V = 5$, Brown's construction and the random construction are comparable. For $V \geq 6$, the random construction beats Brown's construction and gives the best known examples.)

If we consider matrices with no $V \times W$ minor with $V \leq W$, then Kollár, Rónyai, and Szabó [**KRS**] gave examples showing that Theorem 8.9 is sharp whenever $W > V!$.

EXERCISE 8.2. This is a rather long exercise explaining Brown's construction. Recall that the "sphere" $S(x,r) \subset \mathbb{F}_q^3$ is defined as follows:

$$S(x,r) := \{y \in \mathbb{F}_q^3 | \sum_i (x_i - y_i)^2 = r\}.$$

We will consider the $q^3 \times q^3$ matrix with rows indexed by $x \in \mathbb{F}_q^3$ and columns indexed by $y \in \mathbb{F}_q^3$, and with a 1 in the (x,y) position if and only if $y \in S(x,r)$. For well-chosen q and r, we will prove that this matrix has no 3×3 minor of all 1's, and that the number of 1's in the matrix is at least $q^5 - O(q^4)$. We break the argument into six steps.

Step 1. Suppose that q is odd so that 2 is invertible in $\mathbb{F}_q$. Prove that the intersection $S(u,r) \cap S(v,r)$ lies in the plane defined by the equation:

$$\sum_i (u_i - v_i) y_i = \sum_i (u_i - v_i) 2^{-1} (u_i + v_i).$$

In $\mathbb{R}^3$ this plane would be the perpendicular bisector of the segment from u to v. We denote this plane by $\text{Perp}(u, v)$.

Step 2. Given a point $u \in \mathbb{F}_q^3$ and given $\text{Perp}(u, v)$, show how to recover the vector v. In other words, prove that if $\text{Perp}(u, v) = \text{Perp}(u, w)$, then $v = w$.

Step 3. If $u, v, w \in \mathbb{F}_q^3$ are distinct points, show that $S(u, r) \cap S(v, r) \cap S(w, r)$ lies in a line.

Step 4. If the sphere $S(0, r)$ does not contain any line, then show that for any three distinct points $u, v, w \in \mathbb{F}_q^3$, $|S(u, r) \cap S(v, r) \cap S(w, r)| \leq 2$.

Suppose that $S(0, r)$ does not contain any line. Then by Step 4, the incidence matrix between the spheres $\{S(x, r)\}_{x \in \mathbb{F}_q^3}$ and the points $y \in \mathbb{F}_q^3$ has no 3×3 minor of all 1's. The number of 1's in this matrix is $q^3|S(0, r)|$. So to finish Brown's example, we need to choose r so that $S(0, r)$ does not contain a line and $|S(0, r)| \sim q^2$.

We will prove that these two properties hold if q is a prime of the form $4n + 1$ and if r is a quadratic non-residue. Since q is prime of the form $4n + 1$, -1 is a quadratic residue in $\mathbb{F}_q$.

Step 5. Show that the sphere $S(0, r) \subset \mathbb{F}_q^3$ contains a line if and only if r is a quadratic residue.

Hint: Suppose that $S(0, r)$ contains the line parametrized by $\gamma(t) = \vec{m}t + \vec{b}$, where $\vec{m} = (m_1, m_2, m_3)$ and $\vec{b} = (b_1, b_2, b_3)$. This is equivalent to saying that

$$\sum_i (m_i t + b_i)^2 - r = 0 \text{ for all } t \in \mathbb{F}_q.$$

If we expand the left-hand side as a polyomial in t, each coefficient must vanish, leading to the equations

$$\sum_i m_i^2 = 0; \sum_i m_i b_i = 0; \sum b_i^2 = r.$$

These formulas lead to a tricky way of writing r as a negative square. Since -1 is a quadratic residue, we can conclude that r is a quadratic residue.

$$m_1^2 r = m_1^2 \sum b_i^2 = (m_1 b_1)^2 + m_1^2(b_2^2 + b_3^2) =$$

$$= (m_2 b_2 + m_3 b_3)^2 - (m_2^2 + m_3^2)(b_2^2 + b_3^2) = -(m_2 b_3 - m_3 b_2)^2.$$

There is a lot of algebra involved in this derivation, and it would be interesting to understand it in a more conceptual way. It might help a little to consider the following analogous situation in $\mathbb{R}^3$. Consider the hyperboloid in $\mathbb{R}^3$ defined by $x_1^2 + x_2^2 - x_3^2 = r$. If $r > 0$ (i.e. if r is a square), then the hyperboloid is connected and it contains infinitely many lines. On the other hand, if $r < 0$, (i.e. if r is a non-square), then the hyperboloid has two sheets, and it contains no lines. We can see that it has no lines geometrically as follows. We have the equation $x_3^3 = x_1^2 + x_2^2 - r > 0$, and so the hyperboloid does not intersect the plane $x_3 = 0$. But then the only line that could possibly lie in the hyperboloid must be tangent to the (x_1, x_2)-plane. But on such a line, x_3 and r are fixed, and $x_1^2 + x_2^2$ is unbounded, and so such a line does not lie in the hyperboloid either. This is a geometric argument explaining which hyperboloids contain lines. There is also a purely algebraic argument, similar to the argument in Step 5.

Step 6. Suppose that q is a prime of the form $4n + 1$. Prove the following formula for the number of points in $S(0, r) \subset \mathbb{F}_q^3$. If r is zero, $|S(0, r)| = q^2$. If r is a non-zero quadratic residue, then $|S(0, r)| = q^2 + q$. If r is a quadratic non-residue,

then $|S(0,r)| = q^2 - q$. Hints: First check that $|S(0,r_1)| = |S(0,r_2)|$ if r_1, r_2 are any two quadratic non-residues. Similarly, $|S(0,r_1)| = |S(0,r_2)|$ if r_1, r_2 are both (non-zero) quadratic residues. Next, when $r = 0$ or $r = s^2$, count the number of points in $S(0,r)$ by hand.

8.4. Reguli

A regulus is a tool from classical algebraic geometry for studying lines in $\mathbb{R}^3$. Chazelle, Edelsbrunner, Guibas, Pollack, Seidel, Sharir, and Snoeyink used reguli in [**CEGPSSS**] to study the joints problem. Reguli play an important role in studying the incidence geometry of lines in $\mathbb{R}^3$.

One example of a regulus is the surface $z = xy$. We saw this surface in Chapter 3. Each point in the surface $z = xy$ lies in two lines in the surface. Choosing L lines in the surface, we found an example of L lines with $\sim L^2$ intersection points, even though no three of the lines lie in a plane.

The theory of reguli that we present here works over any field with more than two elements. For the rest of the section, we let $\mathbb{F}$ denote a field with more than two elements.

In this section, we will define reguli, and learn their properties and how to use them. Here is a fundamental result about lines in $\mathbb{F}^3$ that leads to the theory of reguli: any three lines in $\mathbb{F}^3$ lie in the zero set of a degree 2 polynomial.

PROPOSITION 8.14. For any three lines l_1, l_2, l_3 in $\mathbb{F}^3$, there is a non-zero degree 2 polynomial Q that vanishes on all three lines.

PROOF. We will prove this result by counting dimensions. We can think of the argument as an example of the polynomial method.

Let $\text{Poly}_2(\mathbb{F}^3)$ be the space of polynomials of degree ≤ 2 in three variables. The space $\text{Poly}_2(\mathbb{F}^3)$ is a vector space of dimension 10. (A basis is given by $x^2, xy, xz, y^2, yz,\ z^2, x, y, z, 1$.)

We choose three points on each line. Let $p_{i,j}$ be three distinct points on l_i. (At this step, we used that $\mathbb{F}$ contains at least three elements.) We have a total of nine points. By linear algebra, we can find a non-zero degree 2 polynomial $Q \in \text{Poly}_2(\mathbb{F}^3)$ that vanishes at all the points $p_{i,j}$. Since Q has degree 2 and vanishes at three distinct points of l_i, it must vanish on all of l_i. So Q vanishes on all three lines as desired. □

This proposition allows us get good information about the lines that intersect all three lines l_1, l_2, and l_3. Exactly what happens depends a little on the properties of l_1, l_2, and l_3. Recall that two lines in $\mathbb{F}^3$ are skew if they don't intersect and they're not parallel. The most important case concerns three skew lines.

PROPOSITION 8.15. If l_1, l_2, and l_3 are pairwise skew, then there is an irreducible degree 2 algebraic surface $R(l_1, l_2, l_3)$ which contains every line that intersects l_1, l_2, and l_3.

PROOF. By the last proposition, there is a non-zero degree 2 polynomial Q that vanishes on l_1, l_2, and l_3. Let $R(l_1, l_2, l_3)$ be the zero set of Q. Suppose that l intersects l_1, l_2, and l_3. Since l_1, l_2, and l_3 are disjoint, the line l must intersect R in three distinct points. But then Q vanishes identically on l, and l is contained in R.

Finally, if Q was reducible, then it would be a product of linear factors, and R would be a union of two planes. But since the lines l_1, l_2, and l_3 are skew, no two of them lie in a plane, and so R cannot be a union of two planes. Also, if Q had degree 1, then R would be a plane, and this cannot happen either. □

The surface $R(l_1, l_2, l_3)$ is called a regulus. We define a regulus to be any irreducible degree 2 surface in $\mathbb{F}^3$ that contains three pairwise-skew lines.

To complement our understanding of three skew lines, we record a couple of trivial lemmas which deal with the case when two lines are not skew.

LEMMA 8.16. Suppose that l_1 and l_2 are lines in $\mathbb{F}^3$ that intersect at a point p. Suppose that P is the plane that contains l_1 and l_2. Then any line which intersects both l_1 and l_2 either contains p or lies in P.

LEMMA 8.17. Suppose that l_1 and l_2 are parallel. Let P be the plane that contains them. Then any line which intersects both l_1 and l_2 lies in P.

The paper [**CEGPSSS**] applied these results to incidence geometry of lines in $\mathbb{F}^3$. For example, they proved that the number of joints determined by L lines in $\mathbb{F}^3$ is $\lesssim L^{7/4}$. We will demonstrate their method by considering the following question: if $\mathfrak{L}$ is a set of L lines in $\mathbb{F}^3$ with ≤ 10 lines in any plane or degree 2 surface, how many intersection points can $\mathfrak{L}$ have? We first met this question in Chapter 3. We will study it with several different methods during the book. Here is our first result on the problem.

THEOREM 8.18. Suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{F}^3$ with ≤ 10 lines in any plane or degree 2 surface. Then the number of intersection points of $\mathfrak{L}$ is $\lesssim L^{5/3}$.

Later in the book, we will return to this question. When $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$, we will prove that the number of intersection points of $\mathfrak{L}$ is $\lesssim L^{3/2}$.

PROOF. A simple intersection point of $\mathfrak{L}$ is a point that lies in exactly two lines of $\mathfrak{L}$. First we bound the number of simple intersection points of $\mathfrak{L}$. Then we refine our analysis to bound the total number of intersection points.

Let us define a square matrix M with rows and columns corresponding to the lines of $\mathfrak{L}$. The matrix M has a 1 in the entry corresponding to row l_i and column l_j if l_i and l_j intersect in a simple intersection point. Otherwise, the entry of M is zero. (By convention, the diagonal entries of M are zero.) The number of simple intersection points is half the number of 1's in the matrix M.

We claim that M has no 3×10 minor of all 1's. Then by Theorem 8.9, it follows that M has $\lesssim L^{5/3}$ 1's, and so $\mathfrak{L}$ has $\lesssim L^{5/3}$ simple intersection points. Suppose for contradiction that M has a 3×10 minor of all 1's. Let l_1, l_2, l_3 be the three lines corresponding to the rows in this 3×10 minor, and let $\tilde{l}_j$ be the lines corresponding to the 10 columns in this minor. First suppose that the lines l_1, l_2, l_3 are pairwise skew. In this case, all the 10 lines $\tilde{l}_j$ lie in the regulus $R(l_1, l_2, l_3)$, contradicting our hypothesis. Otherwise, two of the lines l_1, l_2, l_3 are coplanar. By relabelling, suppose that l_1 and l_2 are coplanar. By hypothesis, each line $\tilde{l}_j$ intersects both l_1 and l_2 at a simple intersection point. Therefore, the 10 lines $\tilde{l}_j$ also lie in the plane containing l_1 and l_2, contradicting our hypothesis.

Now we make an analysis of the higher multiplicity intersection points. We let A be the intersection matrix of $\mathfrak{L}$. In other words, the rows and columns of A are indexed by $\mathfrak{L}$, and the entry a_{ij} is 1 if and only if l_i and l_j intersect. We make the

convention that the diagonal entries of A are zero. We let A_t be the matrix with a 1 in the (i,j)-entry if l_i and l_j intersect at a point lying in $\sim 2^t$ lines of $\mathfrak{L}$. (More precisely, $\sim 2^t$ means $> 2^{t-1}$ and $\leq 2^t$.) The number of points with intersection multiplicity $\sim 2^t$ is $\sim |A_t| 2^{-2t}$. Therefore, the number of intersection points is

$$\sim \sum_{t \geq 1} |A_t| 2^{-2t}.$$

Our next goal is to estimate $|A_t|$.

LEMMA 8.19. Suppose that $\mathfrak{L}$ has ≤ 10 lines in any plane or degree 2 surface. Then A_t has no $3 \times 20 \cdot 2^t$ minor of all 1's.

PROOF. Suppose that A_t has a $3 \times 20 \cdot 2^t$ minor of all 1's. Let the three rows by labelled by l_1, l_2, l_3, and let the columns be labelled by $\tilde{l}_j$ for $j = 1, ..., 20 \cdot 2^t$. If l_1, l_2, l_3 are all skew, then each line $\tilde{l}_j$ lies in the degree 2 surface $R(l_1, l_2, l_3)$. This contradicts our hypothesis. Suppose that l_1, l_2, l_3 are not all skew. After relabelling, we can assume that l_1 and l_2 are not skew. If l_1 and l_2 interesect in a point p and lie in a plane P, then we either get $10 \cdot 2^t$ column lines containing p or $10 \cdot 2^t$ column lines lying in P. By the definition of A_t, there should only be $\leq 2^t$ lines of $\mathfrak{L}$ containing p. And by hypothesis, there are ≤ 10 lines in any plane. So we get another contradiction. Finally, if l_1 and l_2 are parallel lines in the plane P, then we get $20 \cdot 2^t$ lines in the plane P, another contradiction. Since all the cases lead to a contradiction, we see that A_t has no $3 \times 20 \cdot 2^t$ minor of all 1's. □

Knowing that the matrix A_t does not have any $3 \times 20 \cdot 2^t$ minors of all 1's controls the number of 1's in the matrix by the Kővári-Sós-Turán Theorem, Theorem 8.9 above. We recall the statement in our case:

Suppose that A is an $L \times L$ matrix whose entries are 0 or 1. Suppose that A has no $V \times W$ minor of all 1's, for some integers $V \leq W$. Then the number of 1's in A is at most $C(V) W^{1/V} L^{\frac{2V-1}{V}}$.

Plugging in the last lemma, we see that $|A_t| \lesssim 2^{t/3} L^{5/3}$. This gives the following bound on the total number of intersection points of $\mathfrak{L}$:

$$\sum_t |A_t| 2^{-2t} \lesssim L^{5/3} \sum_t 2^{-(5/3)t} \lesssim L^{5/3}.$$

□

EXERCISE 8.3. If l_1, l_2, l_3 are three skew lines in $\mathbb{F}^3$, prove that there is only one degree 2 algebraic surface containing them. In other words, there is only one regulus $R(l_1, l_2, l_3)$.

EXERCISE 8.4. Suppose that $R(l_1, l_2, l_3) \subset \mathbb{R}^3$ is a regulus in $\mathbb{R}^3$. Show that R is a smooth surface. Show that each point of R lies in at most two lines in R. What happens over other fields?

EXERCISE 8.5. (The joints estimate from [**CEGPSSS**].) We describe the main idea of the joints estimate from [**CEGPSSS**], using reguli.

If $\mathfrak{L}$ has ≤ 10 lines in any plane or degree 2 surface, then we know by Theorem 8.18 that $\mathfrak{L}$ has $\lesssim L^{5/3}$ intersection points, and in particular $\lesssim L^{5/3}$ joints. At the other extreme, the lines of $\mathfrak{L}$ can all lie in a plane or a regulus. If all lines of $\mathfrak{L}$ lie in a plane, then $\mathfrak{L}$ has no joints. Using the previous exercise, show that if all the

lines of $\mathfrak{L}$ lie in a regulus R, then $\mathfrak{L}$ has no joints. We need a method to deal with all the inbetween cases.

First, generalize Lemma 8.18 to get a bound for the number of intersection points of a set of L lines in $\mathbb{R}^3$ with at most B lines in any plane or regulus. Then, using this bound, prove by induction that a set of L lines determines at most $CL^{7/4}$ joints.

Here is an outline of the induction. The reader will have to choose a judicious value of B. If $\mathfrak{L}$ contains at most B lines in any plane or regulus, then we use the bound mentioned in the last paragraph. If Σ is a plane or regulus containing more than B lines of $\mathfrak{L}$, then we break up $\mathfrak{L}$ into $\mathfrak{L}_\Sigma$ and $\mathfrak{L}'$, where $\mathfrak{L}_\Sigma$ are the lines in Σ and $\mathfrak{L}'$ are the other lines. As we remarked above, $\mathfrak{L}_\Sigma$ has no joints. The number of joints of $\mathfrak{L}$ involving some lines from $\mathfrak{L}'$ and some lines from $\mathfrak{L}_\Sigma$ is bounded by $2L$, because each line of $\mathfrak{L}'$ intersects Σ in at most two points. Therefore,

$$|J(\mathfrak{L})| \leq 2L + |J(\mathfrak{L}')|, |\mathfrak{L}'| < L - B.$$

Using induction, we can bound $|J(\mathfrak{L}')| \leq C(L-B)^{7/4}$. For a judicious choice of B, the induction will close.

https://doi.org/10.1090//ulect/064/09

CHAPTER 9

Partial symmetries

In [**ElSh**] Elekes and Sharir introduced a very different approach to the distinct distance problem based on partial symmetries.

Suppose G is a group acting on a space X. If $P \subset X$ is a finite set, then we can look at the symmetries of P under the group action. We define

$$G(P) := \{g \in G \text{ such that } g(P) = P\}.$$

A partial symmetry of P is a group element that maps a large chunk of P to another large chunk of P. More precisely we define the r-rich partial symmetries by

$$G_r(P) := \{g \in G \text{ such that } |g(P) \cap P| \geq r\}.$$

The set $G_r(P) \subset G$ is not a subgroup. Perhaps because it lacks this algebraic structure, it hasn't been studied until recently. Elekes started the study of partial symmetries. One important question is to understand the maximum possible size of $G_r(P)$ in different situations.

The group of rigid motions of the plane is a symmetry group for the distinct distance problem: if g is a rigid motion, and P is a set of points, then $d(P) = d(g(P))$. But before [**ElSh**], it was not clear that this symmetry group was an important feature of the problem. In their approach, the symmetries - or partial symmetries - of P play a central role.

In this chapter, we introduce partial symmetries and give some examples. Then we describe a sequence of connections (due to [**ElSh**]), beginning with the distinct distance problem, going through partial symmetries, and ending with the incidence geometry of lines in $\mathbb{R}^3$.

We will ultimately prove the following estimate about the distinct distance problem.

THEOREM 9.1. If $P \subset \mathbb{R}^2$ is a set of N points, then P determines $\gtrsim N(\log N)^{-1}$ distinct distances.

In this chapter, we will use partial symmetries to connect this theorem to the incidence geometry of lines in $\mathbb{R}^3$. We will prove that Theorem 9.1 follows from Theorems 8.3 and 8.4 about lines in $\mathbb{R}^3$.

9.1. Partial symmetries of sets in the plane

Let G be the group of orientation-preserving rigid motions of the plane. Suppose that $P \subset \mathbb{R}^2$ is a finite set. The r-rich partial symmetries of P are defined as follows:

$$G_r(P) := \{g \in G \text{ such that } |g(P) \cap P| \geq r\}.$$

We will study how big $G_r(P)$ can be in terms of r and $|P|$.

For a generic set of N points, $|G_r(P)| = 1$ for $r \geq 3$ and $|G_2(P)| = \binom{N}{2} + 1$. (The number $\binom{N}{2} + 1$ comes up as follows: for each pair of points in P, there is a unique $g \in G$ that switches the two points in the pair. The 2-rich rigid motions are these $\binom{N}{2}$ transpositions and the identity.)

The most interesting example is a square grid of points. If P is a square grid of N points, then $|G_r(P)| \sim N^3 r^{-2}$ for all $2 \leq r \leq N/2$. I have found it surprisingly hard to give a clean proof of this estimate. We are interested in this result mainly to build intuition and to practice thinking about partial symmetries. Therefore, we will give here a heuristic (non-rigorous) argument why $|G_r(P)| \sim N^3 r^{-2}$. We will outline a rigorous proof in the exercises at the end of the chapter.

We can suppose that P is the grid of integer points (x_1, x_2) with $|x_1|, |x_2| \leq M$, and with $N \sim M^2$. We begin by considering translations, which are easy to analyze. If g is a translation by an integer vector (v_1, v_2) with $|(v_1, v_2)| \leq (1/4)M$, then $|g(P) \cap P| \geq N/2$. There are $\sim N$ such choices for v, and this proves that $G_r(P) \gtrsim N$ for every $r \leq N/2$.

Next we consider (orientation-preserving) rotations. For a rotation ρ, we define $\Lambda(\rho) := \rho^{-1}(\mathbb{Z}^2) \cap \mathbb{Z}^2$. Understanding $\Lambda(\rho)$ will help up to understand $\rho(P) \cap P$. This $\Lambda(\rho)$ is always a subgroup of $\mathbb{Z}^2$. Also, if $\Lambda(\rho)$ is non-zero, then it has a nice structure. Suppose that v is a minimum-length (non-zero) vector in $\Lambda(\rho)$. Besides the rotation ρ, we also want to consider a rotation by angle $\pi/2$, which we denote by $\rho_{\pi/2}$. We claim that $\rho_{\pi/2}(v)$ is also in $\Lambda(\rho)$. This happens because $\rho_{\pi/2}$ is an isomorphism of $\mathbb{Z}^2$, and so $\rho_{\pi/2}(v)$ and $\rho_{\pi/2}(\rho(v)) = \rho(\rho_{\pi/2}(v))$ are both integer points. Second, we claim that $\Lambda(\rho)$ is equal to the span of v and $\rho_{\pi/2}(v)$ – otherwise, $\Lambda(\rho)$ would contain a non-zero vector shorter than v. Finally, we see that $\Lambda(\rho)$ has only four minimal vectors: $\pm v$ and $\pm\rho_{\pi/2}(v)$.

Let $|\Lambda(\rho)|$ denote the length of a minimal vector in $\Lambda(\rho)$. The cardinality of $\rho^{-1}(P) \cap P$ is $\sim (M/|\Lambda(\rho)|)^2$. Next we would like to estimate the number of rotations ρ with $|\Lambda(\rho)| \sim S$. This estimate is the most important and trickiest point in the discussion. We will give a non-rigorous argument suggesting that the number of such rotations is roughly S^2. If $|\Lambda(\rho)| \sim S$, then there is an integer vector v with $|v| \sim S$ so that $\rho(v) \in \mathbb{Z}^2$. Once we know v and $\rho(v)$, we have determined ρ. There are $\sim S^2$ possible v in $\mathbb{Z}^2$ with $|v| \sim S$. For each v, the possible choices for $\rho(v)$ are integer vectors that lie on the circle $|x| = |v|$. There are always at least eight such vectors: if $v = (v_1, v_2)$, then the vectors $(\pm v_1, \pm v_2)$ and $(\pm v_2, \pm v_1)$ all qualify. Four of these choices lead to ρ being a rotation by a multiple of $\pi/2$. For these rotations, v is not minimal (unless v was a unit vector). If we choose $\rho(v) = (v_1, -v_2)$, then there seems to be a good chance that v is a minimal vector for ρ (as v_1 and v_2 have no common factor). This discussion suggests that there is often at least one choice of $\rho(v)$ so that the resulting rotation ρ has v as a minimal vector. So as a heuristic we may expect that there are $\gtrsim S^2$ rotations ρ with $|\Lambda(\rho)| \sim S$. On the other hand, for any $\varepsilon > 0$, the circle $|x| = |v|$ contains $\lesssim_\varepsilon S^\varepsilon$ integer points on it. Therefore, the number of rotations ρ with $|\Lambda(\rho)| \sim S$ is $\lesssim_\varepsilon S^{2+\varepsilon}$.

Now we can count rigid motions. Suppose that g is a rigid motion taking p to q, where $p, q \in P$. We write τ_v for the translation by the vector v. There is a unique rotation ρ_g so that we can write g as a composition:

$$g = \tau_q \circ \rho_g \circ \tau_{-p}.$$

We are interested in $|g(P) \cap P|$. We always have $|g(P) \cap P| \lesssim M^2 |\Lambda(\rho_g)|^{-2}$, and this upper bound is sharp as long as p and q are not too close to the edge of P. If we want $|g(P) \cap P| \sim r$, then we must have $|\Lambda(\rho_g)| \sim M r^{-1/2}$, which gives us $\sim N r^{-1}$ choices of ρ_g.

The rest of the discussion is a standard double-counting argument. We consider all the choices $p, q \in P$ and ρ_g with $|\Lambda(\rho_g)| \sim M r^{-1/2}$. We have $\sim N^3 r^{-1}$ choices. For each choice, $\tau_q \circ \rho_g \circ \tau_{-p}$ belongs to $G_r(P)$. But we have overcounted: we counted each element of $G_r(P)$ roughly r times, once for each point $q \in g(P) \cap P$. Therefore, the number of elements of $G_r(P)$ is roughly $N^3 r^{-2}$.

We will come back to counting the partial symmetries of the square grid in an exercise at the end of the chapter.

The paper [**ElSh**] conjectured that the grid example is optimal up to constant factors. They proved the result for $r = 3$, and the general conjecture was proven in [**GK2**].

THEOREM 9.2. Let $P \subset \mathbb{R}^2$ be a set of N points. For any $r \geq 2$,

$$|G_r(P)| \lesssim N^3 r^{-2}.$$

In the next section, we will see that Theorem 9.2 implies Theorem 9.1 about the distinct distance problem. In the following few sections we will see that Theorem 9.2 connects to the incidence geometry of lines in $\mathbb{R}^3$. In particular, we will see that Theorem 9.2 follows from Theorems 8.3 and 8.4, our main results about the incidence geometry of lines in $\mathbb{R}^3$.

9.2. Distinct distances and partial symmetries

Let P be a finite set in the plane $\mathbb{R}^2$. If the distance set $d(P)$ is small, then we will see that P must have lots of partial symmetries.

If $d(P)$ is small, then it must often happen that the same distance occurs between various pairs of points. We can capture this by talking about the distance quadruples, $Q(P)$. They are defined as follows:

$$Q(P) := \{(p_1, q_1, p_2, q_2) \in P^4 \text{ such that } |p_1 - q_1| = |p_2 - q_2| \neq 0\}.$$

For example, if P is a generic set, then the only distance quadruples have the form (p, q, q, p) or (p, q, p, q) for $p, q \in P$ with $p \neq q$. In this case, the number of distance quadruples is $4\binom{N}{2}$.

If there are few distinct distances, then it sounds reasonable that $Q(P)$ must be large, and we make this precise in the following lemma.

LEMMA 9.3. For any set of N points $P \subset \mathbb{R}^2$, the following holds:

$$|d(P)||Q(P)| \geq (N^2 - N)^2.$$

PROOF. Let the distances in $d(P)$ be $d_1, ..., d_{|d(P)|}$. Recall that $d(P)$ is defined to be the set of distances $|p - q|$ among pairs of *distinct* points $p, q \in P$, so the distances d_j are non-zero. Let n_j be the number of ordered pairs $(p, q) \in P^2$ with $|p - q| = d_j$. Then we have

$$Q(P) = \sum_{j=1}^{|d(P)|} n_j^2.$$

On the other hand $\sum_{j=1}^{|d(P)|} n_j = N^2 - N$ is just the number of ordered pairs $(p, q) \in P^2$ with $p \neq q$. Now using Cauchy-Schwarz, we see

$$N^2 - N = \sum_{j=1}^{|d(P)|} (n_j \cdot 1) \leq \left(\sum_j n_j^2 \right)^{1/2} \left(\sum_{j=1}^{|d(P)|} 1 \right)^{1/2} = |Q(P)|^{1/2} |d(P)|^{1/2}.$$

□

For example, if P is a generic set, we see $d(P) = (1/2)(N^2 - N)$ and $Q(P) = 2(N^2 - N)$, so we get equality in the lemma.

A more interesting observation is that $Q(P)$ is closely connected to the number of partial symmetries of P. We state this in the following Proposition.

PROPOSITION 9.4. Let $P \subset \mathbb{R}^2$. Then

$$Q(P) = \sum_{r \geq 2} (2r - 2) |G_r(P)| \sim \sum_{r \geq 2} r |G_r(P)|.$$

PROOF. There is a natural map from $Q(P)$ to $G_2(P)$, which comes from the following lemma.

LEMMA 9.5. Suppose that $(p_1, q_1, p_2, q_2) \in Q(P)$. In other words, we have $(p_1, q_1, p_2, q_2) \in P^4$ and $|p_1 - q_1| = |p_2 - q_2| \neq 0$. Then there is a unique $g \in G$ so that $g(p_1) = p_2$ and $g(q_1) = q_2$.

PROOF. The set of $g \in G$ taking p_1 to p_2 is obtained by applying the translation by $p_2 - p_1$, followed by a rotation around p_2. Since $|p_1 - q_1| = |p_2 - q_2| \neq 0$, there is exactly one such rotation so that the map takes q_1 to q_2. □

We define the map $E : Q(P) \to G_2(P)$ as follows. For any quadruple (p_1, q_1, p_2, q_2) in $Q(P)$, we define $E(p_1, q_1, p_2, q_2)$ to be the unique $g \in G$ so that $g(p_1) = p_2$ and $g(q_1) = q_2$. The letter E stands for Elekes, who first defined this map. We will use the map E to help count $Q(P)$. It's important to note that the map E is not injective. Instead, we have the following lemma.

LEMMA 9.6. Suppose that $g \in G$ and $|g(P) \cap P| = r$. Then $|E^{-1}(g)| = 2\binom{r}{2}$.

PROOF. The set $E^{-1}(g) \subset Q(P)$ is the set of distance quadruples of the form $(p_1, q_1, g(p_1), g(q_1))$. The pair $(g(p_1), g(q_1))$ must lie in $g(P) \cap P$. We get one quadruple of $E^{-1}(g)$ for each ordered pair of distinct elements in $g(P) \cap P$. There are $2\binom{r}{2}$ such ordered pairs. □

If we let $|G_{=r}(P)|$ be the set $\{g \in G \text{ such that } |g(P) \cap P| = r\}$, then we get

$$Q(P) = \sum_{r=2}^{|P|} 2\binom{r}{2} |G_{=r}(P)|.$$

We can rewrite this sum in terms of $|G_r(P)|$ by using the fact that $|G_{=r}(P)| = |G_r(P)| - |G_{r+1}(P)|$.

$$Q(P) = \sum_{r=2}^{|P|} 2\binom{r}{2} |G_{=r}(P)| = \sum_{r \geq 2} 2\binom{r}{2} \left(|G_r(P)| - |G_{r+1}(P)| \right) =$$

$$= \sum_{r \geq 2} |G_r(P)| \left(2\binom{r}{2} - 2\binom{r-1}{2} \right) = \sum_{r \geq 2} (2r - 2) |G_r(P)|.$$

This finishes the proof of Proposition 9.4. □

In this section, we have studied the relationship between $|G_r(P)|$, $|Q(P)|$, and $|d(P)|$. Using Theorem 9.2, we get the following estimates for $|Q(P)|$ and $|d(P)|$:

COROLLARY 9.7. If $P \subset \mathbb{R}^2$ is a set of N points, then $|Q(P)| \lesssim N^3 \log N$.

PROOF. By Proposition 9.4, we have $|Q(P)| \sim \sum_{r=2}^{N} r|G_r(P)|$. By Theorem 9.2,

$$|Q(P)| \sim \sum_{r=2}^{N} r|G_r(P)| \lesssim \sum_{r=2}^{N} N^3 r^{-1} \sim N^3 \log N.$$

□

We can now prove Theorem 9.1: if $P \subset \mathbb{R}^2$ is a set of N points, then $|d(P)| \gtrsim N(\log N)^{-1}$.

PROOF. By Lemma 9.3, we know that $|d(P)||Q(P)| \gtrsim N^4$. By Corollary 9.7 we know that $|Q(P)| \lesssim N^3 \log N$. Therefore, we see that $|d(P)| \gtrsim N(\log N)^{-1}$.

□

We finish this section by returning to the example of a square grid. For a square grid of N points, as we discussed above, $|G_r(P)| \sim N^3 r^{-2}$ for all $2 \leq r \leq N/2$. Therefore, $|Q(P)| \sim N^3 \log N$. On the other hand, for the square grid, $|d(P)| \sim N(\log N)^{-1/2}$, not $\sim N(\log N)^{-1}$. We lost the factor $(\log N)^{1/2}$ when we applied the Cauchy-Schwarz inequality inside the proof of Lemma 9.3. The Cauchy-Schwarz inequality is sharp if each distance $d_i \in d(P)$ occurs the same number of times. In the case of a square grid, different distances occur with different frequencies.

9.3. Incidence geometry of curves in the group of rigid motions

We have seen Theorem 9.1, an estimate about distinct distances, follows from Theorem 9.2, an estimate about partial symmetries. But this problem of partial symmetries sounds hard. It's not at all clear how to get started. The paper [**ElSh**] next describes how partial symmetries are connected to an incidence geometry problem about a natural class of curves in the group G.

For any $p_1, p_2 \in \mathbb{R}^2$, define

$$S_{p_1,p_2} := \{g \in G \text{ such that } g(p_1) = p_2\}.$$

If $p_1 = p_2$ then S_{p_1,p_2} is a subgroup of G. If $p_1 \neq p_2$, then S_{p_1,p_2} is a coset of a subgroup. In any case, S_{p_1,p_2} is a 1-dimensional smooth curve in G, diffeomorphic to a circle. If $P \subset \mathbb{R}^2$, then some of the geometry of P is encoded in the set of curves $\{S_{p_1,p_2}\}_{p_1,p_2 \in P}$. In particular, these curves are connected to the partial symmetries of P by the following Proposition.

PROPOSITION 9.8. If $P \subset \mathbb{R}^2$, then $G_r(P)$ is exactly the set of $g \in G$ that lie in $\geq r$ of the curves $\{S_{p_1,p_2}\}_{p_1,p_2 \in P}$.

PROOF. First suppose that $g \in G_r(P)$. By definition, there is a subset $A_1 \subset P$ with $|A_1| = r$ and $A_2 = g(A_1) \subset P$. Let the points of A_1 be $\{p_{j,1}\}$ for $j = 1, ..., r$. Let $p_{j,2} = g(p_{j,1}) \in A_2 \subset P$. For each $j = 1, ..., r$, we see that $g \in S_{p_{j,1},p_{j,2}}$. So g lies in $\geq r$ of the curves $\{S_{p_1,p_2}\}_{p_1,p_2 \in P}$.

On the other hand, suppose that g lies in $\geq r$ of the curves $\{S_{p_1,p_2}\}_{p_1,p_2 \in P}$. In particular, suppose that $g \in S_{p_{j,1},p_{j,2}}$ for $j = 1, ..., r$, where the pairs $(p_{j,1}, p_{j,2})$

are all distinct. We claim that the points $p_{j,1}$ are all distinct. If $p_{j,1} = p_{j',1}$, then $p_{j,2} = g(p_{j,1}) = g(p_{j',1}) = p_{j',2}$, and then the pairs $(p_{j,1}, p_{j,2})$ and $(p_{j',1}, p_{j',2})$ would be the same. Define $A_1 := \cup_{j=1}^{r} p_{j,1} \subset P$. We see that $A_1 \subset P$ with $|A_1| = r$, and $g(A_1) = \cup p_{j,2} \subset P$. Therefore, $g \in G_r(P)$. □

Figure 9.1 contains two pictures illustrating the same element $g \in G_3(P)$. The first picture takes place in the plane $\mathbb{R}^2$ and it illustrates the sets A_1 and A_2. The second picture takes place in the group G, and it illustrates the element g and the curves $S_{p_{j,1},p_{j,2}}$.

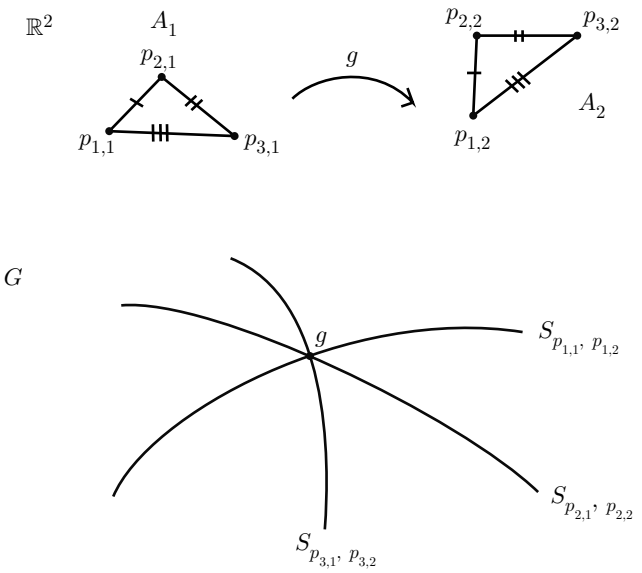

FIGURE 9.1. A 3-rich partial symmetry g.

Because of Proposition 9.8, estimating $|G_r(P)|$ is equivalent to estimating the number of r-rich points of the curves $\{S_{p_1,p_2}\}_{p_1,p_2 \in P} \subset G$. This is a problem of incidence geometry involving curves in a 3-dimensional space. It is similar in spirit to problems about lines in $\mathbb{R}^3$, and we will see in the next section that it is equivalent to a problem about lines in $\mathbb{R}^3$.

As we remarked above, the curves S_{p_1,p_2} are cosets of 1-parameter subgroups of G. In the future, it may be an interesting direction for incidence geometry to work directly in Lie groups. Instead of working in $\mathbb{R}^n$ and studying the incidence geometry of k-planes or k spheres, one can work in a Lie group G and study the cosets of a subgroup $H \subset G$. In this book, we don't know how to exploit this coset structure in a useful way. Instead we change coordinates to reduce to a problem about lines in $\mathbb{R}^3$.

9.4. Straightening coordinates on G

In this section, we define some useful coordinates on most of the group G. In these coordinates, the curves S_{p_1,p_2} become straight lines. We let $G^{trans} \subset G$ be the translations, and we let $G' := G \setminus G^{trans}$. The translations make up only a small part of G, and it's easy to bound $|G_r(P) \cap G^{trans}|$. We will choose coordinates

$\rho : G' \to \mathbb{R}^3$ so that the image of each curve S_{p_1,p_2} is a straight line in $\mathbb{R}^3$. In this way, estimating $|G_r(P) \cap G'|$ reduces to a problem in incidence geometry about lines in $\mathbb{R}^3$.

First we dispense with G^{trans}.

LEMMA 9.9. If $P \subset \mathbb{R}^2$ is a set of N points, then for any $r \geq 2$, $|G_r(P) \cap G^{trans}| \lesssim N^3 r^{-2}$.

PROOF. We first count the quadruples in $E^{-1}(G^{trans}) \subset Q(P)$. Each such quadruple has the form (p_1, q_1, p_2, q_2) with $g(p_1) = p_2$ and $g(q_1) = q_2$ for some translation g. Because g is a translation, we must have $q_2 - q_1 = p_2 - p_1$. Now the number of such quadruples in P^4 is $\leq N^3$, because q_2 is determined by the other three variables.

By Lemma 9.6, each element of $G_r(P) \cap G^{trans}$ has $2\binom{r}{2}$ preimages in $E^{-1}(G^{trans})$. Therefore,

$$|G_r(P) \cap G^{trans}| \leq N^3 (2\binom{r}{2})^{-1} \lesssim N^3 r^{-2}.$$

□

Next we define the coordinates $\rho : G' \to \mathbb{R}^3$. If $g \in G$ is not a translation, then it must be a rotation around a fixed point (x, y) by an angle $\theta \in (0, 2\pi)$. The functions x, y and θ define coordinates on G'. Our coordinates ρ are a small variation of these. In terms of x, y, and θ, we can define ρ by the formula

$$\rho(g) = (x, y, \cot(\theta/2)).$$

It's straightforward to check that ρ is a bijection: we note that $\theta/2 \in (0, \pi)$ and that the cotangent function is a bijection from $(0, \pi)$ to $\mathbb{R}$.

The point of this definition lies in the following lemma.

LEMMA 9.10. For any $p_1, p_2 \in \mathbb{R}^2$, $\rho(S_{p_1,p_2} \cap G')$ is a straight line l_{p_1,p_2} in $\mathbb{R}^3$. If $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$, then this straight line can be parametrized as follows. Let a be the midpoint of p_1 and p_2: $a = (1/2)(p_1 + p_2)$. Let v be the vector formed by rotating $(1/2)(p_2 - p_1)$ by $\pi/2$. In equations $v = (1/2)(y_1 - y_2, x_2 - x_1)$. Now $\rho(S_{p_1,p_2} \cap G')$ is equal to the line parametrized by $t \mapsto (a + tv, t)$.

PROOF. Suppose that $g \in G'$ with $g(p_1) = p_2$, and that g is a rotation around (x, y) by angle θ. The distance from (x, y) to p_1 must be the same as the distance from (x, y) to p_2, and so (x, y) must lie on the perpendicular bisector of p_1 and p_2. Next, we consider how the angle θ depends on the center (x, y). Consider Figure 9.2.

The points (x, y), a and p_2 form a right triangle, with right angle at a. The angle of this triangle at (x, y) is $\theta/2$. The vector from a to p_2 is $(1/2)(p_2 - p_1)$. Recall that v is the vector formed by rotating $(1/2)(p_2 - p_1)$ by $\pi/2$. So the vector from a to (x, y) is in the direction of v, and it has length $\cot(\theta/2)|v|$. Therefore, $(x, y) = a + \cot(\theta/2)v$.

This formula shows how (x, y) depends on θ. If we define $t = \cot(\theta/2)$, then $(x, y) = a + tv$. Since $\rho(x, y, \theta) = (x, y, \cot(\theta/2))$, we see that $\rho(S_{p_1,p_2} \cap G')$ is equal to the line parametrized by $t \mapsto (a + tv, t)$.

(For completeness, we should also draw a second picture. See Figure 9.3

In this slightly funny picture, p_1 is a small positive rotation from p_2. Since g takes p_1 to p_2, and since we defined θ to lie in $(0, 2\pi)$, θ is actually $> \pi$ in

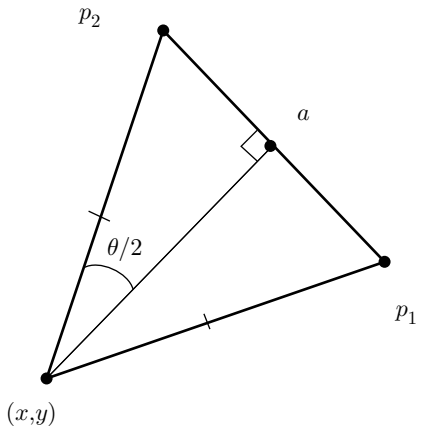

FIGURE 9.2

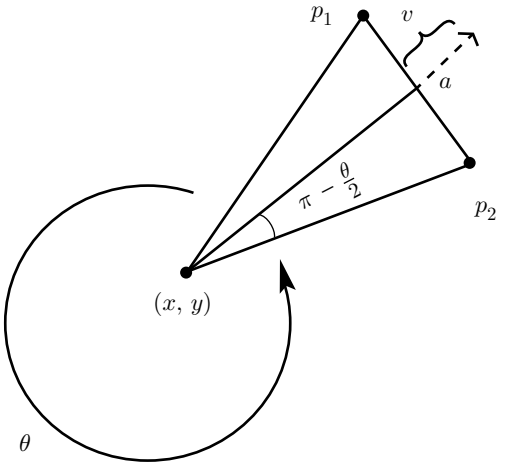

FIGURE 9.3. This figure illustrates the case $\theta > \pi$.

this picture. It is still true that $(x, y), a$, and p_2 make a right triangle with a right angle at a. In this case, the angle of the triangle at (x, y) is $\pi - \theta/2$. Note that $\cot(\theta/2) = -\cot(\pi - \theta/2)$. On the other hand, in this picture, the vector v points in the opposite direction of the vector from a to (x, y). Therefore, $(x, y) = a + \cot(\pi - \theta/2)(-v) = a + \cot(\theta/2)v$.)

□

These coordinates are quite useful because lines are easier to understand than other curves. I don't have any more general perspective on this construction, and it may just be a fortuitous coincidence. Using these coordinates, problems about partial symmetries of sets in $\mathbb{R}^2$ can be translated into problems about the incidence geometry of lines in $\mathbb{R}^3$.

9.5. Applying incidence geometry of lines to partial symmetries

Let P be a set of N points in $\mathbb{R}^2$. Let $\mathfrak{L}(P)$ be the set of lines $\{l_{p_1,p_2}\}_{p_1,p_2\in P}$. We recall from Lemma 9.10 that $l_{p_1,p_2} = \rho(S_{p_1,p_2}\cap G')$ and that l_{p_1,p_2} is parametrized by $t \mapsto (a+tv, t)$ where a, v are described as follows. If $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$, then

$$a = \frac{1}{2}(x_1 + x_2, y_1 + y_2); v = \frac{1}{2}(y_1 - y_2, x_2 - x_1). \tag{9.1}$$

First we check that the lines l_{p_1,p_2} are all distinct.

LEMMA 9.11. For each $a, v \in \mathbb{R}^2$, there is a unique $p_1, p_2 \in \mathbb{R}^2$ obeying equation 9.1. Therefore, the lines $\{l_{p_1,p_2}\}_{p_1,p_2\in\mathbb{R}^2}$ are all distinct.

PROOF. Given a, v, we can find $x_1 + x_2$ and $x_2 - x_1$, and then we can find x_1 and x_2. Similarly, we can find $y_1 + y_2$ and $y_1 - y_2$, and then we can find y_1 and y_2. If $(a, v) \neq (a', v')$, then the lines parametrized by $t \mapsto (at+v, t)$ and $t \mapsto (a't+v', t)$ are distinct. Therefore, the lines $\{l_{p_1,p_2}\}_{p_1,p_2\in\mathbb{R}^2}$ are all distinct. □

In particular, we see that $\mathfrak{L}(P)$ is a set of N^2 straight lines in $\mathbb{R}^3$. The r-rich points of $\mathfrak{L}(P)$ correspond to r-rich partial symmetries of P. More precisely, we have the following lemma:

LEMMA 9.12. $|G_r(P) \cap G'| = |P_r(\mathfrak{L}(P))|$.

PROOF. By Proposition 9.8, $G_r(P) \cap G'$ is exactly the set of $g \in G'$ that lie in $\geq r$ of the curves $\{S_{p_1,p_2}\}_{p_1,p_2\in P}$. Since $\rho : G' \to \mathbb{R}^3$ is a bijection, the result follows. □

We see that the partial symmetries of P are related to the incidence structure of $\mathfrak{L}(P)$. Our goal is to prove Theorem 9.2, saying that $|G_r(P)| \lesssim N^3 r^{-2}$. So we would like to prove that the number of r-rich points determined by $\mathfrak{L}(P)$ is $\lesssim N^3 r^{-2} = |\mathfrak{L}(P)|^{3/2} r^{-2}$ for each $2 \leq r \leq N$. We may first ask if these inequalities holds for any set of N^2 lines in $\mathbb{R}^3$. This inequality can fail badly if the lines cluster in a plane or a degree 2 surface. For any $2 \leq r \leq N$, the grid construction from Section 7.1 gives a configuration of N^2 lines in a plane with $\sim N^4 r^{-3}$ r-rich points. Also, the regulus construction from Section 3.5 gives a configuration of N^2 lines in a regulus with $\sim N^4$ 2-rich points.

This leads us to ask how many lines of $\mathfrak{L}(P)$ can lie in a plane or a degree 2 surface.

LEMMA 9.13. For any degree $D \geq 1$, there is a constant $C(D)$ so that the following holds. If P is a set of N points in the plane, then $\mathfrak{L}(P)$ contains $\leq C(D)N$ lines in any degree D algebraic surface.

We will prove Lemma 9.13 in Section 9.6 below.

We can now prove Theorem 9.2 using Lemma 9.13 and our main results on the incidence geometry of lines in $\mathbb{R}^3$: Theorems 8.3 and 8.4. These theorems immediately give the following estimate:

THEOREM 9.14. If $\mathfrak{L}$ is a set of N^2 lines in $\mathbb{R}^3$ with $\leq N$ lines in any plane or degree 2 surface, and if $2 \leq r \leq N$, then the number of r-rich points of $\mathfrak{L}$ is $\lesssim N^3 r^{-2}$.

We want to apply Theorem 9.14 to $\mathfrak{L}(P)$, but there is a tiny wrinkle, because the hypothesis of Theorem 9.14 requires that $\mathfrak{L}$ contains $\leq N$ lines in any plane or degree 2 surface, while Lemma 9.13 tells us that $\mathfrak{L}(P)$ contains at most CN lines in any plane or degree 2 surface. To get around this tiny wrinkle, we enlarge $\mathfrak{L}$ to a set of $(N')^2$ lines $\mathfrak{L}'$, with $N' \lesssim N$ so that $\mathfrak{L}'$ contains at most N' lines in any plane or regulus. Then Theorem 9.14 implies that $|P_r(\mathfrak{L}(P))| \leq |P_r(\mathfrak{L}')| \lesssim (N')^3 r^{-2} \lesssim N^3 r^{-2}$. This proves that $|G_r(P) \cap G'| = |P_r(\mathfrak{L}(P))| \lesssim N^3 r^{-2}$. By Lemma 9.9, we know that $|G_r(P) \cap G^{trans}| \lesssim N^3 r^{-2}$. So all together, $|G_r(P)| \lesssim N^3 r^{-2}$, establishing Theorem 9.2.

Except for checking Lemma 9.13, we have now proven that Theorems 8.3 and 8.4 imply Theorem 9.2, our estimate for partial symmetries.

We will develop several tools for studying the incidence geometry of lines in $\mathbb{R}^3$. The first important tool is the polynomial partitioning method, which we study in Chapter 10. Using just polynomial partitioning, we will prove a slightly weaker incidence estimate for lines in $\mathbb{R}^3$:

THEOREM 9.15. For any $\varepsilon > 0$, there are constants $D(\varepsilon)$ and $C(\varepsilon)$ so that the following holds. If $\mathfrak{L}$ is a set of N^2 lines in $\mathbb{R}^3$ with at most N lines in any algebraic surface of degree $\leq D(\varepsilon)$, and $2 \leq r \leq N$, then

$$|P_r(\mathfrak{L})| \leq C(\varepsilon) N^{3+\varepsilon} r^{-2}.$$

The proof of Theorem 9.15 is significantly shorter than the proof of Theorem 9.14. Theorem 9.15 leads to slightly weaker estimates about partial symmetries, distance quadruples, and distinct distances. By the same arguments as above, Theorem 9.15 implies that if $P \subset \mathbb{R}^2$ is a set of N points, then $|G_r(P)| \lesssim_\varepsilon N^{3+\varepsilon} r^{-2}$, $|Q(P)| \lesssim_\varepsilon N^{3+\varepsilon}$ and $|d(P)| \gtrsim_\varepsilon N^{1-\varepsilon}$.

This finishes our discussion of the connection between distinct distances, distance quadruples, partial symmetries, and the incidence geometry of lines in $\mathbb{R}^3$.

9.6. The lines of $\mathfrak{L}(P)$ don't cluster in a low degree surface

Now we come back to the proof of Lemma 9.13, which says that the lines of $\mathfrak{L}(P)$ cannot cluster in a low degree surface. Let us recall the statement.

LEMMA. For any degree $D \geq 1$, there is a constant $C(D)$ so that the following holds. If P is a set of N points in the plane, then $\mathfrak{L}(P)$ contains $\lesssim_D N$ lines in any degree D algebraic surface.

PROOF. We begin with the case of a plane (a degree 1 surface). We state this as its own lemma.

LEMMA 9.16. If $P \subset \mathbb{R}^2$ is a set of N points, then there are at most N lines of $\mathfrak{L}(P)$ in any plane.

PROOF. Let l_{p_1,p_2} be the line $\rho(G' \cap S_{p_1,p_2})$. For a fixed $p_1 \in \mathbb{R}^2$, any two lines l_{p_1,p_2} and $l_{p_1,p_2'}$ are skew. The curves S_{p_1,p_2} and $S_{p_1,p_2'}$ are disjoint because g cannot map p_1 to both p_2 and p_2'. Therefore, the lines l_{p_1,p_2} and $l_{p_1,p_2'}$ are also disjoint. Next we check that these lines are not parallel. To do that, we recall the parametrization for l_{p_1,p_2} given in Lemma 9.10: $t \mapsto (a + tv, t)$, where a is the midpoint $(1/2)(p_1 + p_2)$, and where v is the rotation by $\pi/2$ of $(1/2)(p_2 - p_1)$. The vector $(v, 1)$ is parallel to l_{p_1,p_2}. Similarly, the vector $(v', 1)$ is parallel to $l_{p_1,p_2'}$,

where v' is the rotation by $\pi/2$ of $(1/2)(p_2' - p_1)$. We have $v' \neq v$, and so l_{p_1,p_2} and $l_{p_1,p_2'}$ are not parallel.

For a fixed $p_1 \in \mathbb{R}^2$, a plane contains at most one of the lines $\{l_{p_1,p_2}\}_{p_2 \in \mathbb{R}^2}$. Hence a plane can contain at most N of the lines $\{l_{p_1,p_2}\}_{p_1,p_2 \in P}$. □

Now we turn to the higher degree case which will take more work. Suppose that Q is an irreducible polynomial of degree at most D and that Q is not degree 1. It suffices to prove that $Z(Q)$ contains $\leq 3D^2N$ lines of $\mathfrak{L}(P)$.

Fix $p \in P$ and consider the set of lines $\{l_{p,p'}\}_{p' \in P}$. It can happen that all N of these lines lie in a regulus. This occurs if P is contained in a circle or in a line; we will discuss the examples more in the next section. But we will see that if $Z(Q)$ is not a plane, then there is at most one point p so that $Z(Q)$ contains many lines $\{l_{p,q}\}_{q \in \mathbb{R}^2}$. We let $\mathfrak{L}_p := \{l_{p,q}\}_{q \in \mathbb{R}^2}$.

LEMMA 9.17. If Q is an irreducible polynomial of degree at most D and Q is not degree 1, then there is at most one point $p \in \mathbb{R}^2$ so that $Z(Q)$ contains at least $2D^2$ lines of $\mathfrak{L}_p$.

Given Lemma 9.17, the proof of Lemma 9.13 is straightforward. For $N-1$ of the points $p \in P$, $Z(Q)$ contains at most $2D^2$ of the lines $\{l_{p,p'}\}_{p' \in P}$. For the last point $p \in P$, $Z(Q)$ contains at most all N of the lines $\{l_{p,p'}\}_{p' \in P}$. In total, $Z(Q)$ contains at most $(2D^2+1)N$ lines of $\mathfrak{L}(P)$.

The proof of Lemma 9.17 is based on a more technical lemma which describes the algebraic structure of the set of lines $\{l_{p,q}\}$ in $\mathbb{R}^3$.

LEMMA 9.18. For each p, each point of $\mathbb{R}^3$ lies in a unique line from the set $\{l_{p,q}\}_{q \in \mathbb{R}^2}$. Moreover, for each p, there is a non-vanishing vector field $V_p(x_1, x_2, x_3)$, so that at each point, $V_p(x)$ is tangent to the unique line $l_{p,q}$ through x. Moreover, $V_p(x)$ is a polynomial in p and x, with degree at most 1 in the p variables and degree at most 2 in the x variables.

Let us assume this technical lemma for the moment and use it to prove Lemma 9.17.

Fix a point $p \in \mathbb{R}^2$. Suppose $Z(Q)$ contains at least $2D^2$ lines from the set $\mathfrak{L}_p := \{l_{p,q}\}_{p,q \in \mathbb{R}^2}$. On each of these lines, Q vanishes identically, and V_p is tangent to the line. Therefore, $V_p \cdot \nabla Q$ vanishes on all these lines. But $V_p \cdot \nabla Q$ is a polynomial in x of degree at most $2D-2$. If $V_p \cdot \nabla Q$ and Q have no common factor, then a version of Bezout's theorem, Theorem 6.7, implies that there are at most $2D^2 - 2D$ lines where the two polynomials vanish. Therefore, $V_p \cdot \nabla Q$ and Q have a common factor. Since Q is irreducible, Q must divide $V_p \cdot \nabla Q$, and we see that $V_p \cdot \nabla Q$ vanishes identically on $Z(Q)$.

Now suppose that $Z(P)$ contains at least $2D^2$ lines from $\mathfrak{L}_{p_1}$ and from $\mathfrak{L}_{p_2}$. We see that $V_{p_1} \cdot \nabla Q$ and $V_{p_2} \cdot \nabla Q$ vanish on $Z(Q)$. For each fixed x, the expression $V_p \cdot \nabla Q$ is a degree 1 polynomial in p. Therefore, for any point p in the affine span of p_1 and p_2, $V_p \cdot \nabla Q$ vanishes on $Z(Q)$.

Suppose that $Z(Q)$ has a non-singular point x, which means that $\nabla Q(x) \neq 0$. In this case, x has a smooth neighborhood $U_x \subset Z(Q)$ where ∇Q is non-zero. If $V_p \cdot \nabla Q$ vanishes on $Z(Q)$, then the vector field V_p is a vector field on U_x, and so its integral curves lie in U_x. But the integral curves of V_p are exactly the lines of $\mathfrak{L}_p$. Therefore, for each p on the line connecting p_1 and p_2, the line of $\mathfrak{L}_p$ through x lies in $Z(Q)$. Since x is a smooth point, all of these lines must lie in the tangent plane

$T_xZ(Q)$, and we see that $Z(Q)$ contains infinitely many lines in a plane. Using Bezout's theorem, Theorem 6.7, again, we see that $Z(Q)$ is a plane, and that Q is a degree 1 polynomial. This contradicts our assumption that $\operatorname{Deg} Q > 1$.

We have now proven Lemma 9.17 in the case that $Z(Q)$ contains a non-singular point. But if every point of $Z(Q)$ is singular, then we get an even stronger estimate on the lines in $Z(Q)$:

LEMMA 9.19. Suppose that Q is a non-zero irreducible polynomial of degree D on $\mathbb{R}^3$. If $Z(Q)$ has no non-singular point, then $Z(Q)$ contains at most D^2 lines.

PROOF. Since every point of $Z(Q)$ is singular, ∇Q vanishes on $Z(Q)$. In particular, each partial derivative $\partial_i Q$ vanishes on $Z(Q)$. We suppose that $Z(Q)$ contains more than D^2 lines and derive a contradiction. Since $\partial_i Q = 0$ on $Z(Q)$ and $Z(Q)$ contains more than D^2 lines, then Bezout's theorem, Theorem 6.7, implies that Q and $\partial_i Q$ have a common factor. Since Q is irreducible, Q must divide $\partial_i Q$. Since $\operatorname{Deg} \partial_i Q < \operatorname{Deg} Q$, it follows that $\partial_i Q$ is identically zero for each i. This implies that Q is constant. By assumption, Q is not the zero polynomial and so $Z(Q)$ is empty. But we assumed that $Z(Q)$ contains at least D^2+1 lines, giving a contradiction. □

This finishes the proof of Lemma 9.17 assuming Lemma 9.18. It only remains to prove Lemma 9.18.

First we check that each point $x \in \mathbb{R}^3$ lies in exactly one of the lines $\{l_{p,q}\}_{q\in\mathbb{R}^2}$. Suppose $p = (p_1, p_2)$ and $q = (q_1, q_2)$ are points in $\mathbb{R}^2$. By Lemma 9.10, x lies in $l_{p,q}$ if and only if the following equation holds for some $t \in \mathbb{R}$.

$$\left(\frac{p_1+q_1}{2}, \frac{p_2+q_2}{2}, 0\right) + t\left(\frac{p_2-q_2}{2}, \frac{q_1-p_1}{2}, 1\right) = (x_1, x_2, x_3).$$

Given p and x, we can uniquely solve this equation for t and $q = (q_1, q_2)$. First of all, we see that $t = x_3$. Next we get a matrix equation for q_1, q_2 of the following form:

$$\begin{pmatrix} 1 & -x_3 \\ x_3 & 1 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = a_p(x),$$

where $a_p(x)$ is a vector whose entries are polynomials in x, p of degree ≤ 1 in x and degree ≤ 1 in p. Since the determinant of the matrix on the left-hand side is $1 + x_3^2 > 0$, we can solve this equation for q_1 and q_2. The solution has the form

$$q_1 = (x_3^2+1)^{-1} b_{1,p}(x); q_2 = (x_3^2+1)^{-1} b_{2,p}(x), \tag{9.2}$$

where b_1, b_2 are polynomials in x, p of degree ≤ 2 in x and degree ≤ 1 in p.

We have now proven that each point of $\mathbb{R}^3$ lies in a unique line from the set $\{l_{p,q}\}_{q\in\mathbb{R}^2}$. Now we can construct the vector field V_p. From Lemma 9.10, we see that the vector $(p_2 - q_2, q_1 - p_1, 2)$ is tangent to $l_{p,q}$. If $x \in l_{p,q}$, then we can use Equation 9.2 to expand q in terms of x, p, and we see that the following vector field is tangent to $l_{p,q}$ at x:

$$v_p(x) := (p_2 - (x_3^2+1)^{-1} b_{2,p}(x), (x_3^2+1)^{-1} b_{1,p}(x) - p_1, 2).$$

The coefficients of $v_p(x)$ are not polynomials because of the $(x_3^2+1)^{-1}$. We define $V_p(x) = (x_3^2+1)v_p(x)$, so

$$V_p(x) = \left(p_2(x_3^2+1) - b_{2,p}(x), b_{1,p}(x) - p_1(x_3^2+1), 2x_3^2+2\right).$$

The vector field $V_p(x)$ is tangent to the family of lines $\{l_{p,q}\}_{q\in\mathbb{R}^2}$. Moreover, V_p never vanishes because its last component is $2x_3^2 + 2$. Therefore, the integral

curves of V_p are exactly the lines $\{l_{p,q}\}_{q\in\mathbb{R}^2}$. Moreover, each component of V_p is a polynomial of degree ≤ 2 in x and degree ≤ 1 in p.

This finishes the proof of Lemma 9.18 and hence the proof of Lemma 9.13. □

9.7. Examples of partial symmetries related to planes and reguli

We have been exploring the connection between the geometry of a finite set $P \subset \mathbb{R}^2$ and the incidence properties of the associated lines $\mathfrak{L}(P)$ in $\mathbb{R}^3$. Problems about partial symmetries $G_r(P)$ or distance quadruples $Q(P)$ can be seen from two perspectives: the perspective of $\mathbb{R}^2$ or the perspective of the lines $\mathfrak{L}(P)$ in $\mathbb{R}^3$. In this section, we consider some examples which might help give a feel for how these two perspectives are related. We won't need these examples anywhere in the book; they are just for context. In particular, we focus on examples involving planes and reguli in $\mathbb{R}^3$ – what do planes and reguli in $\mathbb{R}^3$ correspond to in $\mathbb{R}^2$? We will give one example in detail and suggest other examples as exercises.

For $q \in \mathbb{R}^2$ and $r > 0$, let $S^1(q, r)$ denote the circle around q of radius r. Let $S(p, S^1(q, r))$ be the set of $g \in G'$ so that $g(p) \in S^1(q, r)$. We use the ρ coordinates on G', as defined in Section 9.4, so we can think of $S(p, S^1(q, r)) \subset \mathbb{R}^3$.

Proposition 9.20. For any $p, q \in \mathbb{R}^2$ and $r > 0$, $S(p, S^1(q, r))$ lies in a regulus.

Proof. It's easiest to describe the rulings of the regulus first. For any $q' \in S^1(q, r)$, $l_{p,q'}$ is in $S(p, S^1(q, r))$. This gives one ruling. Now observe that $g(p) \in S^1(q, r)$ if and only if $g^{-1}(q) \in S^1(p, r)$. For $p' \in S^1(p, r)$, $l_{p',q}$ lies in $S(p, S^1(q, r))$. These lines give the other ruling.

Now we prove that $S(p, S^1(q, r))$ lies in a regulus. We pick three points q'_1, q'_2, q'_3 in $S^1(q, r)$, and consider the three lines $l_i = l_{p,q'_i}$. As we saw above, these three lines all lie $S(p, S^1(q, r))$. From the formula for the line $l_{p,q}$, it is easy to check that any two of the lines $\{l_{p,q}\}_{q\in\mathbb{R}^2}$ are skew – we used this fact in the proof of Lemma 9.16. In particular, the lines l_1, l_2, l_3 are skew, and there is a regulus $R(l_1, l_2, l_3)$ that contains them. If l is any other line that intersects l_1, l_2, and l_3, then $l \subset R(l_1, l_2, l_3)$ as well.

If $p' \in S(p, r)$, then for each q'_i, there is a unique $g \in G$ so that $g(p) = q'_i$ and $G(p') = q$. Therefore, for any point $p' \in S^1(p, r)$, the curve $S_{p',q}$ intersects each $S(p, q'_i)$. For almost every $p' \in S^1(p, r)$, $l_{p',q}$ intersects each l_i. Therefore, for almost every $p' \in S(p, r)$, $l_{p',q}$ lies in $R(l_1, l_2, l_3)$.

If we replace q'_1, q'_2, q'_3 by other points in $S^1(q, r)$, we get a new regulus that still contains almost every line $l_{p',q}$. Since these two reguli intersect in infinitely many lines, they must actually be the same regulus. Therefore, $R(l_1, l_2, l_3)$ contains $l_{p,q'}$ for every $q' \in S^1(q, r)$. But $S(p, S^1(q, r)) = \cup_{q'\in S^1(q,r)} l_{p,q'} \subset R(l_1, l_2, l_3)$. □

(Remark. With some more work and more algebraic geometry, I think that it's possible to show that $S(p, S^1(q, r))$ is equal to the regulus R described above.)

We sketch some more examples of this flavor as exercises for the reader.

Exercise 9.1. Suppose that $\lambda \subset \mathbb{R}^2$ is a line. (In this section, we use l for lines in $\mathbb{R}^3$ and λ for lines in $\mathbb{R}^2$.) We define $S(p, \lambda)$ to be the set of $g \in G'$ so that $g(p) \in \lambda$. We use the ρ coordinates on G', so we can think of $S(p, \lambda)$ as a subset of $\mathbb{R}^3$. Prove that $S(p, \lambda)$ is contained in a regulus.

One approach is to find the two rulings of the regulus as in the last Proposition. We give some hints about this appraoch. For any $q \in \lambda$, we see that $l_{p,q} \subset S(p,\lambda)$. These lines give one ruling of the regulus. The other ruling requires a small detour.

Let l be a line in $\mathbb{R}^3$. We call l "horizontal" if it lies in a plane of the form $z = h$. The lines $l_{p,q}$ are not horizontal, as we can see from their parametrization in Lemma 9.10. The horizontal lines also have a nice interpretation in terms of the group of rigid motions. Let λ_1, λ_2 be oriented lines in $\mathbb{R}^2$. Let $S(\lambda_1, \lambda_2)$ be the set of $g \in G'$ so that g maps λ_1 onto λ_2 preserving the orientation.

EXERCISE 9.2. Prove that $S(\lambda_1, \lambda_2)$ is a horizontal line in $\mathbb{R}^3$ and that each horizontal line in $\mathbb{R}^3$ is $S(\lambda_1, \lambda_2)$ for a pair of oriented lines (λ_1, λ_2).

Prove that each non-horizontal line $l \subset \mathbb{R}^3$ is $l_{p,q}$ for a unique pair $p, q \in \mathbb{R}^2$.

Now we can describe the other ruling of $S(p,\lambda)$. For every λ' containing p, $S(\lambda', \lambda)$ lies in $S(p,\lambda)$. These lines give the other ruling of $S(p,\lambda)$.

There is another example of reguli that appears in connection with a distance problem for two lines. Here is the distance problem. Let P be a set of N points on the x axis. Let Q be a set of N points on the y axis. Consider $d(P,Q)$, the set of distances $\{d(p,q)\}_{p\in P, q\in Q}$. For generic P, Q, $|d(P,Q)| = N^2$, but there is a clever choice where $|d(P,Q)| \sim N$. Namely, let $P = \{(\sqrt{a}, 0)\}_{a=1,\dots,N}$, and let $Q = \{(0, \sqrt{b})\}_{b=1,\dots,N}$. The distance from $(\sqrt{a}, 0)$ to $(0, \sqrt{b})$ is $(a+b)^{1/2}$. Since there are $2N-1$ values of $a+b$, $|d(P,Q)| = 2N-1$. Because there are few distinct distances, there are many ($\sim N^3$) quadruples (p_1, q_1, p_2, q_2) with $|p_1 - q_1| = |p_2 - q_2|$, $p_i \in P$, $q_i \in Q$. We can study these quadruples using the lines $\mathfrak{L}(P,Q) := \{l_{p,q}\}_{p\in P, q\in Q}$.

EXERCISE 9.3. Prove that the N^2 lines of $\mathfrak{L}(P,Q)$ are clumped into $\sim N$ reguli with $\sim N$ lines in each regulus. In each regulus, there are $\sim N^2$ intersection points, for a total of $\sim N^3$ intersection points, corresponding to the $\sim N^3$ distance quadruples.

More generally, one can study the distances between points on two lines. Suppose that $P \subset l \subset \mathbb{R}^2$ and $P' \subset l' \subset \mathbb{R}^2$. We define $d(P,P') := \{|p - p'|\}_{p\in P, p'\in P'}$. If l, l' are not parallel or perpendicular, then there are interesting lower bounds for $d(P,P')$. If $|P| = |P'| = N$, then the best known lower bound is $|d(P,P')| \gtrsim N^{4/3}$, due to Sharir, Sheffer and Solymosi [**SSS**].

EXERCISE 9.4. Consider a plane $\pi \subset \mathbb{R}^3$. Identifying $G' = \mathbb{R}^3$, this plane describes some subset of the group of rigid motions G. Describe this subset of the rigid motions in terms of how they act on the plane.

9.8. Other exercises

EXERCISE 9.5. Let P be a square grid of N points. Find $\mathfrak{L}(P)$ and use it to estimate $G_r(P)$. Show that $\mathfrak{L}(P)$ is the set of lines we considered in Exercise 8.1. Using this example, show that $|G_r(P)| \sim N^3 r^{-2}$ for all $2 \le r \le N/400$.

EXERCISE 9.6. Suppose that $P \subset \mathbb{R}^2$ has N points and $|d(P)| = \varepsilon N$. Prove that P has an r-rich partial symmetry for $r \ge e^{c\varepsilon^{-1}}$ for some $c > 0$.

https://doi.org/10.1090//ulect/064/10

CHAPTER 10

Polynomial partitioning

The last few chapters of the book gave a survey of incidence geometry, explaining some of the main theorems and questions in the field. Now that we have this background, we begin the third part of the book, studying applications of the polynomial method to incidence geometry. There will be several chapters, discussing a few different methods of using polynomials in incidence geometry.

In this chapter, we study the polynomial partitioning method. We introduce the method, and we use it to give a different proof of the Szemerédi-Trotter theorem. Then we turn to estimates about lines in $\mathbb{R}^3$, and we start to study the questions from Section 8.1.

The main result of the chapter is Theorem 9.15, which we restate here in a slightly stronger way:

THEOREM 10.1. For any $\varepsilon > 0$, there are constants $D(\varepsilon)$ and $C(\varepsilon)$ so that the following holds. If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with at most $L^{(1/2)+\varepsilon}$ lines in any algebraic surface of degree $\leq D(\varepsilon)$, then

$$|P_r(\mathfrak{L})| \leq C(\varepsilon) L^{(3/2)+\varepsilon} r^{-2} + 2Lr^{-1}.$$

Chapter 9 explains how this type of estimate is connected with partial symmetries and the distinct distance problem. In particular, as explained in Section 9.5, Theorem 10.1 implies that for any $\varepsilon > 0$, any set of N points in the plane determines at least $c_\varepsilon N^{1-\varepsilon}$ distinct distances.

Polynomial partitioning builds on an older partitioning method in incidence geometry, called the cutting method. We start by describing the cutting method, and then we explain how polynomials come into the picture.

10.1. The cutting method

In order to motivate polynomial partitioning, we begin by discussing the first partitioning method in incidence geometry, called the cutting method. The cutting method is a fundamental approach to incidence geometry problems introduced by Clarkson, Edelsbrunner, Guibas, Sharir, and Welzl in [**CEGSW**]. They used the method to reprove the Szemerédi-Trotter theorem and to prove many new results. One interesting feature is that the cutting method also applies to incidence geometry problems in higher dimensions. For example, [**CEGSW**] estimates the number of r-rich points determined by a set of unit spheres in $\mathbb{R}^3$.

In this section, we will describe some of the key ideas of the cutting method. The goal of the section is to give intuition and background, but not complete proofs. There aren't any results in this section that we will use later. Hopefully, it will give a flavor of the cutting method, which will help to understand polynomial partitioning later in the chapter. If you're interested, you can learn much more about the cutting method in the book [**PS**].

The cutting method is a divide-and-conquer approach. We cut the plane into pieces, estimate the number of incidences in each piece, and add up the contributions of the different pieces. To get a sense of how it works, we outline a proof of the Szemerédi-Trotter theorem using the cutting method. We will consider the version of the Szemerédi-Trotter theorem about incidences between points and lines, Theorem 7.11, which we restate here:

THEOREM. If $\mathcal{S}$ is a set of S points in the plane, and $\mathfrak{L}$ is a set of L lines in the plane, then the number of incidences between $\mathcal{S}$ and $\mathfrak{L}$ is bounded as follows:

$$|I(\mathcal{S}, \mathfrak{L})| \lesssim (S^{2/3}L^{2/3} + S + L).$$

In addition to the lines of $\mathfrak{L}$, let us consider a set of D auxiliary lines. These lines cut the plane into pieces, called cells. More formally, the cells are the connected components of the complement of the D lines. In the cutting method, we estimate the number of incidences in each of these cells and add up the contributions. For a generic choice of D auxiliary lines, there are $\sim D^2$ cells. Crucially, each line of $\mathfrak{L}$ can enter only a small fraction of these cells.

LEMMA 10.2. A line can enter at most $D+1$ of the cells determined by D lines.

PROOF. To go from one cell to another, a line must cross one of the D auxiliary lines. But a given line intersects each of the D auxiliary lines at most once. □

This divide-and-conquer approach works best if we can divide the problem into roughly equal pieces. Since each line enters $\sim D$ of the D^2 cells, an average cell intersects $\sim L/D$ lines of $\mathfrak{L}$. We say that the lines are (roughly) equidistributed if

$$\text{Each cell intersects } \lesssim LD^{-1} \text{ lines of } \mathfrak{L}. \qquad (EquiL)$$

Similarly, we say the points of $\mathcal{S}$ are (roughly) equidistributed if

$$\text{Each cell contains } \lesssim SD^{-2} \text{ points of } \mathcal{S}. \qquad (EquiS)$$

Let us suppose for now that we are able to arrange (EquiL) and (EquiS), and let us sketch how to bound $|I(\mathcal{S}, \mathfrak{L})|$.

We know a bound for $|I(\mathcal{S}, \mathfrak{L})|$ using double counting. Recall that the incidence matrix of $(\mathcal{S}, \mathfrak{L})$ encodes which points lie on which line. It has one row for each point of $\mathcal{S}$ and one column for each line of $\mathfrak{L}$. Given a point $x \in \mathcal{S}$ and a line $l \in \mathfrak{L}$, the corresponding entry of the matrix is 1 if $x \in l$ and 0 otherwise. Because two lines intersect in at most one point, the incidence matrix has no 2×2 minor of all 1's. Therefore, the Kővári-Sós-Turán theorem, Theorem 8.9, has the following corollary about the number of incidences between points and lines (Corollary 8.10):

COROLLARY. If $\mathcal{S}$ is a set of S points and $\mathfrak{L}$ is a set of L lines, then

$$I(\mathcal{S}, \mathfrak{L}) \lesssim SL^{1/2} + L. \qquad (*)$$

Instead of applying this corollary to control the incidences between $\mathcal{S}$ and $\mathfrak{L}$, we apply it to control the number of incidences in each cell. Since each cell has $\lesssim SD^{-2}$ points and $\lesssim LD^{-1}$ lines, the number of incidences in each cell is $\lesssim SL^{1/2}D^{-5/2} + LD^{-1}$. The number of cells is $\lesssim D^2$, and so the total number of incidences in all the cells is

$$\lesssim D^{-1/2}SL^{1/2} + DL.$$

There could also be some incidences on the cell walls – on the union of the D auxiliary lines. For simplicity, let us assume that the auxiliary lines are distinct from the lines of $\mathfrak{L}$ (but this point is not crucial to the argument). Each line of $\mathfrak{L}$ has at most D intersection points with the auxiliary lines, so there are at most DL incidences coming from the cell walls. All together, we get the following bound.

$$|I(\mathcal{S}, \mathfrak{L})| \lesssim D^{-1/2} S L^{1/2} + DL.$$

If we optimize the right-hand side over D, we get $|I(\mathcal{S}, \mathfrak{L})| \lesssim S^{2/3} L^{2/3}$. To summarize, if we can arrange equidistribution, then we recover the Szemerédi-Trotter bound. In this analysis, we assumed both (EquiL) and (EquiS), but with a little more work, it turns out that either one of them suffices.

After I had done these calculations, I thought that I had understood the main idea of the proof of the Szemerédi-Trotter theorem. I initially assumed that it wouldn't be so hard to find D auxiliary lines so that $\mathcal{S}$ and/or $\mathfrak{L}$ are equidistributed. My wrong intuition went something like this. Suppose that we choose D auxiliary lines without thinking too much or somehow at random. There is no particular reason why the points of $\mathcal{S}$ should clump into one of the cells instead of another, and so they will probably be fairly evenly distributed. This intuition was totally wrong.

Here is a different heuristic about the equidistribution problem. Suppose that we attempt to equidistribute the points of $\mathcal{S}$. We are going to choose D auxiliary lines, which means that we have $2D$ real variables at our disposal. There will be $\sim D^2$ cells. Equidistribution involves one condition for each cell. Since there are $\sim D^2$ cells, there are $\sim D^2$ conditions that we want to satisfy. Each condition is an inequality, but it's approximately an equality. In each cell, we would like to have $\lesssim SD^{-2}$ points of $\mathcal{S}$. This SD^{-2} is the average number of points per cell, so in a lot of the cells we will have $\sim SD^{-2}$ points of $\mathcal{S}$. Roughly speaking, we have $\sim D$ variables and we are hoping to solve $\sim D^2$ equations. Without other information, this is a method which sounds unlikely to work.

Here is an example of a set $\mathcal{S}$ which is impossible to equidistribute. Let γ be a strictly convex closed curve such as a circle, and suppose that $\mathcal{S}$ is contained in γ. Each auxiliary line meets γ in at most 2 points. Therefore, the curve γ is divided into at most $2D$ pieces by the auxiliary lines. So the points $\mathcal{S}$ lie in at most $2D$ cells. One of these cells must have $\gtrsim SD^{-1}$ points of $\mathcal{S}$, which is much more than SD^{-2}.

To make the cutting method work, one needs a major additional idea in order to find D auxiliary lines with some equidistribution properties. We give a rough sketch of how [**CEGSW**] approach this problem. In the next sections, we discuss polynomial partitioning, which will give a different approach to this issue.

An important idea from [**CEGSW**] is to choose the D auxiliary lines randomly from the lines of $\mathfrak{L}$. If we do that, then the auxiliary lines interact well with the lines of $\mathfrak{L}$. The auxiliary lines don't quite obey (EquiL), but they do have some useful equidistribution properties.

The D auxiliary lines give a polyhedral decomposition of $\mathbb{R}^2$, where the cells are 2-faces, and there are also edges and vertices. One quantity that is fairly easy to estimate is the maximal number of lines of $\mathfrak{L}$ that intersect any edge. To get a little intuition about this, suppose that we first choose $D/10$ auxiliary lines and cut the plane into cells with them. We still have $9D/10$ more lines to choose, which will cut these cells into smaller cells. Consider an edge of the current decomposition

that intersects KLD^{-1} lines of $\mathfrak{L}$ for some $K > 100$. When we choose the next $D/10$ auxiliary lines, we will on average choose $K/10$ lines that intersect our edge. An edge from the first stage that intersects $> 100LD^{-1}$ lines of $\mathfrak{L}$ is likely to be cut into many smaller edges in the second stage. Heuristically, edges that intersect $> 100LD^{-1}$ lines have a short half-life, and rapidly decay into smaller edges. Filling in some details of this argument, it is not hard to show that with high probability, every edge intersects at most $1000(L \log L)D^{-1}$ lines of $\mathfrak{L}$.

If we somehow knew that every cell had $\lesssim 1$ edges, then it would follow that every cell intersects $\lesssim L \log LD^{-1}$ lines of $\mathfrak{L}$, which is very close to (EquiL). However, there are configurations of lines where this fails badly. Suppose that K is a convex polygon with L sides, and let $\mathfrak{L}$ be the set of lines formed by extending each edge of K to be a line. Suppose that we choose D lines of $\mathfrak{L}$. The original convex polygon K lies in one of the components of the complement of the D lines. This component has D edges and all the other lines enter it.

The paper [**CEGSW**] builds a good cell decomposition by starting with the cell decomposition using random lines of $\mathfrak{L}$, and then refining it with some additional well chosen line segments, used to break up cells with many edges. For more details, we refer to the original paper [**CEGSW**] or the book [**PS**].

One of the interesting features of the cutting method is that it gives results in higher dimensions. Suppose that we have S points and L hypersurfaces in $\mathbb{R}^n$. If we randomly choose D of the hypersurfaces, they will cut $\mathbb{R}^n$ into cells. Because we choose the D auxiliary hypersurfaces randomly, we can hope to prove that the L hypersurfaces are fairly equidistributed among the cells. If we do prove an equidistribution bound, then we can estimate the incidences in each cell and add up the contributions. For example, using this approach, [**CEGSW**] was able to prove an interesting estimate about the incidence geometry of unit spheres in $\mathbb{R}^3$. We will come back to this problem using polynomial partitioning in the exercises later in the chapter.

The cutting method leads to interesting estimates about $(n-1)$-dimensional surfaces in $\mathbb{R}^n$ for any n. But there are significant difficulties trying to apply it to k-dimensional surfaces for $k < n-1$, such as lines in $\mathbb{R}^3$.

10.2. Polynomial partitioning

Now we can explain the main idea of polynomial partitioning. Instead of using D hyperplanes to cut space into cells, we use the zero set of a degree D polynomial. Partitioning with D hyperplanes is a special case that happens when the polynomial is a product of D degree 1 factors. If P is a polynomial of degree D, then a line either lies in $Z(P)$ or else it crosses $Z(P)$ at most D times. So each line intersects at most $D+1$ connected components of $\mathbb{R}^n \setminus Z(P)$ – exactly the same bound as if $Z(P)$ was a union of D hyperplanes. Allowing an arbitrary degree D algebraic surface instead of just a union of D planes greatly increases our flexibility, which makes equidistribution much easier to achieve.

The complement of D hyperplanes in $\mathbb{R}^n$ generically has $\sim D^n$ components. The topology of real algebraic varieties was studied by Oleinik-Petrovski, Milnor, and Thom - (see [**Mi**]). They proved among other things that for any degree D polynomial P, $\mathbb{R}^n \setminus Z(P)$ has at most $\sim D^n$ connected components. The vector space $\mathrm{Poly}_D(\mathbb{R}^n)$ also has dimension $\sim D^n$. If we would like $\mathbb{R}^n \setminus Z(P)$ to have $\sim D^n$ connected components, and we would like the objects we are studying to

be equidistributed among these components, then we have roughly D^n degrees of freedom and we want to satisfy $\sim D^n$ conditions. Although this parameter counting doesn't prove anything, it sounds more plausible that we can choose $P \in \mathrm{Poly}_D(\mathbb{R}^n)$ to arrange even distribution.

Now we can state our main result about the existence of good polynomial partitionings.

THEOREM 10.3. For any dimension n, we can choose $C(n)$ so that the following holds. If X is any finite subset of $\mathbb{R}^n$ and D is any degree, then there is a non-zero polynomial $P \in \mathrm{Poly}_D(\mathbb{R}^n)$ so that $\mathbb{R}^n \setminus Z(P)$ is a disjoint union of $\lesssim D^n$ open sets O_i each containing $\leq C(n)|X|D^{-n}$ points of X.

There is a crucial caveat about this theorem. The theorem does NOT guarantee that the points of X lie in the complement of $Z(P)$. In fact it is possible that $X \subset Z(P)$. There are two extreme cases. If all the points of X lie in the complement of $Z(P)$, then we get optimal equidistribution, and we have a good tool to do a divide-and-conquer argument following ideas from the cutting method. If all the points of X lie in $Z(P)$, then we see that X is contained in a surface of controlled degree, and we can try to use algebraic tools to study X. Generally, X will have some points in $Z(P)$ and some points in the complement, and we study each part of X separately.

The proof of Theorem 10.3 uses ideas from topology, which we describe in the next section. These topological ideas connect to several parts of mathematics. In this chapter, we focus on the connection with incidence geometry. Earlier, Gromov used similar ideas to study a problem in differential geometry. His proof has a lot of parallels with the proof of the finite field Kakeya problem. We will come back to discuss these ideas in Chapter 14.

10.3. Proof of polynomial partitioning

The polynomial partitioning theorem, Theorem 10.3, is based on topology. It is based on the Stone-Tukey ham sandwich theorem, which in turn is based on the Borsuk-Ulam theorem. In this section, we introduce all these results, and we prove that Theorem 10.3 follows from the Borsuk-Ulam theorem. We don't give a proof of the Borsuk-Ulam theorem, but we give references to well-written proofs.

10.3.1. Ham sandwich theorems. We will build our polynomial cell decomposition using a tool from topology, the ham sandwich theorem. In this section, we introduce ham sandwich theorems. Here is the first version of the ham sandwich theorem.

THEOREM 10.4. (Ham sandwich theorem, [**StTu**]) If $U_1, ..., U_n$ are finite volume open sets in $\mathbb{R}^n$, then there is a hyperplane that bisects each set U_i.

(Banach proved this theorem in the 3-dimensional case in the late 30's. Stone and Tukey generalized the proof to higher dimensions.)

For example, if each U_i is a round ball, then the solution is a plane that goes through the center of each ball. If the centers are in general position, there will be exactly one solution. We can get a heuristic sense of the situation by counting parameters. The set of hyperplanes in $\mathbb{R}^n$ is given by n parameters. Heuristically, we might expect that the subset of hyperplanes that bisect U_1 is given by $n-1$

parameters; that the subset of hyperplanes that bisect U_1 and U_2 is given by $n-2$ parameters etc.

Stone and Tukey generalized Banach's proof to higher dimensions. They also realized that the same proof gives a much more general version of the ham sandwich theorem. Now we formulate the Stone-Tukey ham sandwich theorem.

Notice that the planes are exactly the zero sets of degree 1 polynomials (polynomials of the form $a_1x_1 + ... + a_nx_n + b$). We can generalize this setup by allowing other functions, such as higher degree polynomials. Suppose that V is a vector space of functions from $\mathbb{R}^n$ to $\mathbb{R}$. Multiplication by a scalar doesn't change the zero set of a function f, so might say heuristically that the family of zero sets is given by $\operatorname{Dim} V - 1$ parameters. For example, if V is the space of polynomials of degree ≤ 1, then $\operatorname{Dim} V = n+1$, and the dimension of the set of hyperplanes is n. Since we have $\operatorname{Dim} V - 1$ parameters to play with, we might hope to bisect $\operatorname{Dim} V - 1$ sets $U_i \subset \mathbb{R}^n$. Stone and Tukey showed that this heuristic is correct under very mild conditions on the space V.

To state our theorem, we make a little basic notation. For any function $f : \mathbb{R}^n \to \mathbb{R}$, we let $Z(f) := \{x \in \mathbb{R}^n | f(x) = 0\}$. We say that f bisects a finite volume open set U if

$$Vol_n\{x \in U | f(x) > 0\} = Vol_n\{x \in U | f(x) < 0\} = (1/2)Vol_nU.$$

THEOREM 10.5. (General ham sandwich theorem, Stone and Tukey, [**StTu**]) Let V be a vector space of continuous functions on $\mathbb{R}^n$. Let $U_1, ..., U_N \subset \mathbb{R}^n$ be finite volume open sets with $N < \operatorname{Dim} V$. For any function $f \in V \setminus \{0\}$, suppose that $Z(f)$ has Lebesgue measure 0. Then there exists a function $f \in V \setminus \{0\}$ which bisects each set U_i.

The ham sandwich theorem is one corollary, given by taking V to be the degree 1 polynomials. If we consider the space of polynomials with degree $\leq D$, we get the following corollary.

COROLLARY 10.6. Let $U_1, ..., U_N \subset \mathbb{R}^n$ be finite volume open sets. Suppose that $N < \binom{D+n}{n} = \operatorname{Dim} \operatorname{Poly}_D(\mathbb{R}^n)$. Then there is a non-zero $P \in \operatorname{Poly}_D(\mathbb{R}^n)$ that bisects all the sets U_i.

PROOF. In Lemma 2.2, we proved that $\operatorname{Dim} \operatorname{Poly}_D(\mathbb{R}^n) = \binom{D+n}{n}$. It's also easy to check that for a non-zero polynomial P, $Z(P)$ has measure 0. We outlined the proof in Exercise 2.4. Now the conclusion follows from the Stone-Tukey ham sandwich theorem. □

The polynomial ham sandwich theorem is analogous to the more basic polynomial existence lemma which we have been using throughout the course. We recall the lemma here to make the analogy clear.

LEMMA 10.7. (Polynomial existence lemma) If $\mathbb{F}$ is a field and if $p_1, ..., p_N \in \mathbb{F}^n$ are points and $N < \binom{D+n}{n} = \operatorname{Dim} \operatorname{Poly}_D(\mathbb{F}^n)$, then there is a non-zero polynomial of degree $\leq D$ that vanishes at each p_i.

The polynomial existence lemma is analogous to the polynomial ham sandwich theorem. The first is based on linear algebra, and the second is based on topology. The polynomial existence lemma was a basic step in all of our arguments in Chapter 2. Using the polynomial ham sandwich theorem instead gives a new direction to the polynomial method.

10.3.2. The proof of the ham sandwich theorem. The proof of the ham sandwich theorem is based on the Borsuk-Ulam theorem.

THEOREM 10.8. (Borsuk-Ulam) Suppose that $\phi : S^N \to \mathbb{R}^N$ is a continuous map that obeys the antipodal condition $\phi(-x) = -\phi(x)$ for all $x \in S^N$. Then the image of ϕ contains 0.

For a proof of the Borsuk-Ulam theorem, the reader can look at Matousek's book *Using the Borsuk-Ulam theorem* [**Ma**] or in the book *Differential Topology* by Guillemin and Pollack, [**GP**], Chapter 2.6. The book *Using the Borsuk-Ulam theorem* discusses some surprising applications of Borsuk-Ulam to combinatorics.

PROOF OF THE GENERAL HAM SANDWICH THEOREM. For each i from 1 to N, we define $\phi_i : V \setminus \{0\} \to \mathbb{R}$ by

$$\phi_i(F) := Vol\left(\{x \in U_i | F(x) > 0\}\right) - Vol\left(\{x \in U_i | F(x) < 0\}\right).$$

So $\phi_i(F) = 0$ if and only if F bisects U_i. Also, ϕ_i is antipodal, $\phi_i(-F) = -\phi_i(F)$.

We will check below that ϕ_i is a continuous function from $V \setminus \{0\}$ to $\mathbb{R}$. We assemble the ϕ_i into one function $\phi : V \setminus \{0\} \to \mathbb{R}^N$.

We know that $\operatorname{Dim} V > N$, and without loss of generality we can assume that $\operatorname{Dim} V = N + 1$. Now we choose an isomorphism of V with $\mathbb{R}^{N+1}$, and we think of S^N as a subset of V. The map $\phi : S^N \to \mathbb{R}^N$ is antipodal and continuous. By the Borsuk-Ulam theorem, there is a function $F \in S^N \subset V \setminus \{0\}$ so that $\phi(F) = 0$. This function F bisects each U_i.

It only remains to check the technical point that ϕ_i is continuous. We state this fact as a lemma.

CONTINUITY LEMMA. Let V be a finite-dimensional vector space of continuous functions on $\mathbb{R}^n$. Suppose that for each $f \in V \setminus \{0\}$, the set $Z(f)$ has measure 0.

If U is a finite volume open set, then the measure of the set $\{x \in U | f(x) > 0\}$ depends continuously on $f \in V \setminus \{0\}$.

PROOF. The proof is based on measure theory. A good reference for measure theory is the book *Real Analysis* by Stein and Shakarchi, [**StSh**].

Suppose that f is a function in $V \setminus \{0\}$ and $f_n \in V \setminus \{0\}$ with $f_n \to f$ in V. A priori, f_n converges to f in the topology of V. But then it follows that $f_n \to f$ pointwise. Pick any $\epsilon > 0$. We can find a subset $E \subset U$ so that $f_n \to f$ uniformly pointwise on $U \setminus E$, and $m(E) < \epsilon$. (See Theorem 4.4 on page 33 of [**StSh**].)

The set $\{x \in U | f(x) = 0\}$ has measure zero. Also U has finite measure. Therefore, we can choose δ so that the set $\{x \in U$ such that $|f(x)| < \delta\}$ has measure less than ϵ. (See Corollary 3.3 on page 20 of [**StSh**].)

Next we choose n large enough so that $|f_n(x) - f(x)| < \delta$ on $U - E$. Then the measures of $\{x \in U | f_n(x) > 0\}$ and $\{x \in U | f(x) > 0\}$ differ by at most 2ϵ. But ϵ was arbitrary. □

This finishes the proof of the Stone-Tukey ham sandwich theorem. □

10.3.3. A ham sandwich theorem for finite sets. We now adapt the ham sandwich theorem to finite sets of points. Instead of open sets U_i, we will have finite sets S_i. We say that a polynomial P bisects a finite set S if at most half the points in S are in $\{P > 0\}$ and at most half the points in S are in $\{P < 0\}$. Note that P may vanish on some or all of the points of S.

COROLLARY 10.9. Let $S_1, \ldots, S_N$ be finite sets of points in $\mathbb{R}^n$ with $N < \binom{D+n}{n} = \operatorname{Dim} \operatorname{Poly}_D(\mathbb{R}^n)$. Then there is a non-zero polynomial $P \in \operatorname{Poly}_D(\mathbb{R}^n)$ that bisects each set S_i.

Let us give an example now. Suppose that we take two sets S_1 and S_2 in the plane, both lying on the x-axis, with $S_1 \subset [0,1] \times \{0\}$ and $S_2 \subset [2,3] \times \{0\}$. Since $2 < \binom{2+1}{2} = 3$, we should be able to choose a degree 1 polynomial P to bisect both S_1 and S_2. The only option is to choose $P = x_1$ so that $Z(P)$ is the x_1-axis. Any line transverse to the x_1-axis will fail to bisect one of the two sets. Because of this situation, we have to allow P to "bisect" a finite set S in the case that P vanishes on S.

The idea of the proof is to replace the finite sets by finite unions of δ-balls, apply the polynomial ham sandwich theorem, and then take $\delta \to 0$.

PROOF. For each $\delta > 0$, define $U_{i,\delta}$ to be the union of δ-balls centered at the points of S_i. By the polynomial ham sandwich theorem we can find a non-zero polynomial P_δ of degree $\leq D$ that bisects each set $U_{i,\delta}$. By rescaling P_δ, we can assume that $P_\delta \in S^N \subset \operatorname{Poly}_D(\mathbb{R}^n) \setminus \{0\}$.

Since S^N is compact, we can find a sequence $\delta_m \to 0$ so that P_{δ_m} converges to a polynomial $P \in S^N \subset \operatorname{Poly}_D(\mathbb{R}^n) \setminus \{0\}$. Since the coefficients of P_{δ_m} converge to the coefficients of P, P_{δ_m} converges to P uniformly on compact sets.

We claim that P bisects each set S_i. We prove the claim by contradiction. Suppose instead that $P > 0$ on more than half of the points of S_i. (The case $P < 0$ is similar.) Let $S_i^+ \subset S_i$ denote the set of points of S_i where $P > 0$. By choosing ϵ sufficiently small, we can assume that $P > \epsilon$ on the ϵ-ball around each point of S_i^+. Also, we can choose ϵ small enough that the ϵ-balls around the points of S_i are disjoint. Since P_{δ_m} converges to P uniformly on compact sets, we can find m large enough that $P_{\delta_m} > 0$ on the ϵ-ball around each point of S_i^+. By making m large, we can also arrange that $\delta_m < \epsilon$. Therefore, $P_{\delta_m} > 0$ on more than half of U_{i,δ_m}. This contradiction proves that P bisects S_i. □

10.3.4. Cell decompositions. Now we are ready to prove our polynomial partitioning theorem, Theorem 10.3. We restate the theorem here:

THEOREM. For any dimension n, we can choose $C(n)$ so that the following holds. If S is any finite subset of $\mathbb{R}^n$ and D is any degree, then there is a non-zero polynomial $P \in \operatorname{Poly}_D(\mathbb{R}^n)$ so that $\mathbb{R}^n \setminus Z(P)$ is a disjoint union of $\lesssim D^n$ open sets O_i each containing $\leq C(n)|S|D^{-n}$ points of S.

PROOF. We construct our polynomial cell decomposition by repeatedly using Corollary 10.9 - the polynomial ham sandwich theorem for finite sets. To begin, we find a polynomial P_1 of degree 1 that bisects S. We divide $\mathbb{R}^n \setminus Z(P_1)$ into two disjoint open sets according to the sign of P_1. We let S_+ denote the points of S where $P_1 > 0$, and S_- denote the points of S where $P_1 < 0$. The sets S_+ and S_- each contain $\leq |S|/2$ points. Next we find a low degree polynomial P_2 that bisects S_+ and S_-. There are four possible sign conditions on P_1 and P_2, and the subset of S for each sign condition contains $\leq |S|/4$ points. The complement of $Z(P_1P_2)$ is a disjoint union of four open sets determined by the signs of P_1 and P_2, and each of these open sets contains at most $|S|/4$ points of S. We continue in this way to define polynomials P_3, P_4, etc.

The polynomial P_j bisects 2^{j-1} finite sets, determined by the signs of $P_1, ..., P_{j-1}$. By Corollary 10.9, we can find P_j with degree $\leq C(n)2^{j/n}$. The complement of $Z(P_1 \cdot ... \cdot P_j)$ is the disjoint union of 2^j open sets corresponding to the possible signs of $P_1, ..., P_j$, and each of these open sets contains $\leq |S|2^{-j}$ points. We repeat this procedure J times, and we define $P = P_1 \cdot ... \cdot P_J$. The sign conditions of $P_1, ..., P_J$ divide $\mathbb{R}^n \setminus Z(P)$ into 2^J open sets O_i each containing at most $|S|2^{-J}$ points of S. (Some of these open sets may be empty.)

Finally we choose D so that $\mathrm{Deg}(P) \leq D$, which means that $C(n) \sum_{j=0}^{J} 2^{j/n} \leq D$. The sum is a geometric sum, and the last term is comparable to the whole. Therefore, we can arrange that $\mathrm{Deg}\, P \leq D$ for $D \leq C(n)2^{J/n}$. The number of points of S in each O_i is $\leq |S|2^{-J} \leq C(n)|S|D^{-n}$. □

10.4. Using polynomial partitioning

Now we discuss how to apply polynomial partitioning in incidence geometry. We begin by reproving the Szemerédi-Trotter theorem. We follow the argument from the paper [**KMS**], which uses polynomial partitioning to reprove several classical theorems of incidence geometry.

We recall the statement of the Szemerédi-Trotter theorem. We proved the theorem in Chapter 7 using the crossing number lemma - see Theorem 7.11 and Theorem 7.1.

THEOREM 10.10. (Szemerédi-Trotter, [**SzTr**]) If $\mathcal{S}$ is a set of S points in $\mathbb{R}^2$ and $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^2$, then the number of incidences obeys the following bound:

$$I(\mathcal{S}, \mathfrak{L}) \lesssim S^{2/3}L^{2/3} + S + L.$$

We first recall a simple estimate for $|I(\mathcal{S}, \mathfrak{L})|$ using double counting.

LEMMA 10.11. If $\mathcal{S}$ and $\mathfrak{L}$ are as above, then

- $I(\mathcal{S}, \mathfrak{L}) \leq L + S^2$.
- $I(\mathcal{S}, \mathfrak{L}) \leq L^2 + S$.

PROOF. Fix $x \in \mathcal{S}$. Let L_x be the number of lines of $\mathfrak{L}$ that contain x and no other point of $\mathcal{S}$. For each other point $y \in S$, there is at most one line of $\mathfrak{L}$ containing x and y. Therefore, $I(x, \mathfrak{L}) \leq S + L_x$. So $I(\mathcal{S}, \mathfrak{L}) \leq S^2 + \sum_{x \in S} L_x \leq S^2 + L$.

The proof of the other inequality is similar. □

To prove the theorem, we will combine polynomial partitioning with this double counting bound. We will use polynomial partitioning to divide the plane into cells, and we will use the double counting bound to estimate the number of incidences in each cell.

PROOF OF THEOREM 10.10. If $L > S^2$ or $S > L^2$, then the conclusion follows from the counting lemma. Therefore, we can now restrict to the case that

$$S^{1/2} \leq L \leq S^2. \tag{10.1}$$

Let D be a degree that we will choose later. By the polynomial partitioning theorem, Theorem 10.3, we can find a non-zero polynomial P of degree $\leq D$ so that each component of the complement of $Z(P)$ contains $\lesssim SD^{-2}$ points of $\mathcal{S}$. Let O_i be the components of $\mathbb{R}^2 \setminus Z(P)$. For each i, let $\mathcal{S}_i = \mathcal{S} \cap O_i$ and let $\mathfrak{L}_i \subset \mathfrak{L}$ be the set of lines of $\mathfrak{L}$ that intersect O_i. Let $S_i = |\mathcal{S}_i|$ and $L_i = |\mathfrak{L}_i|$.

If a line does not lie in $Z(P)$, then it intersects $Z(P)$ in at most D points, and so each line intersects at most $D+1$ cells. Therefore, $\sum L_i \leq (D+1)L$.

Applying Lemma 10.11 in each cell, we get

$$I(\mathcal{S}_i, \mathfrak{L}_i) \leq L_i + S_i^2.$$

We let $\mathcal{S}_{cell}$ be the union of $\mathcal{S}_i$ - all the points of $\mathcal{S}$ that lie in the interiors of the cells.

$$I(\mathcal{S}_{cell}, \mathfrak{L}) = \sum_i I(\mathcal{S}_i, \mathfrak{L}_i) \leq \sum_i L_i + \sum_i S_i^2 \lesssim LD + SD^{-2} \sum_i S_i \leq LD + S^2 D^{-2}.$$

We let $\mathcal{S} = \mathcal{S}_{cell} \cup \mathcal{S}_{alg}$, where $\mathcal{S}_{alg}$ is the set of points in $Z(P)$. It remains to bound $I(\mathcal{S}_{alg}, \mathfrak{L})$. We divide $\mathfrak{L}$ as $\mathfrak{L}_{alg} \cup \mathfrak{L}_{cell}$, where $\mathfrak{L}_{alg}$ are the lines contained in $Z(P)$ and $\mathfrak{L}_{cell}$ are the other lines. The total number of incidences is bounded by

$$|I(\mathcal{S}, \mathfrak{L})| \leq |I(\mathcal{S}_{cell}, \mathfrak{L})| + |I(\mathcal{S}_{alg}, \mathfrak{L}_{cell})| + |I(\mathcal{S}_{alg}, \mathfrak{L}_{alg})|.$$

Each line of $\mathfrak{L}_{cell}$ has at most D intersection points with $Z(P)$, and so it has at most D incidences with $\mathcal{S}_{alg}$. Therefore $I(\mathcal{S}_{alg}, \mathfrak{L}_{cell}) \leq LD$.

The number of lines in $\mathfrak{L}_{alg}$ is at most D. By Lemma 10.11,

$$|I(\mathcal{S}_{alg}, \mathfrak{L}_{alg})| \leq S + D^2.$$

All together, we see that

$$|I(\mathcal{S}, \mathfrak{L})| \lesssim LD + S^2 D^{-2} + S + D^2.$$

Now we choose D to optimize this bound. We can minimize the sum of $LD + S^2D^{-2}$ by choosing $D \sim S^{2/3}L^{-1/3}$. We need to choose $D \geq 1$ an integer. By Inequality 10.1 above, we know that $L \leq S^2$, and so $S^{2/3}L^{-1/3} \geq 1$. Therefore, we can choose D to be a positive integer of size $\sim S^{2/3}L^{-1/3}$. Inequality 10.1 also tells us that $L \geq S^{1/2}$, and so $D^2 \sim S^{4/3}L^{-2/3} \leq S$. Therefore,

$$|I(\mathcal{S}, \mathfrak{L})| \lesssim S^{2/3}L^{2/3} + S.$$

□

The proof of the Szemerédi-Trotter theorem uses the topology of $\mathbb{R}^2$. In this proof, using polynomial partitioning, topology enters twice. The idea of a cell decomposition uses topology. More precisely, we are using topology when we refer to the connected components of $\mathbb{R}^2 \setminus Z(P)$ and when we say that a line enters at most $1 + \operatorname{Deg} P$ cells. Then the proof of the polynomial partitioning theorem uses topology again. It involves the ham sandwich theorem which in turn follows from the Borsuk-Ulam theorem.

10.5. Exercises

In these exercises, we apply polynomial partitioning to various problems in incidence geometry. Instead of lines in the plane, we consider various curves in the plane and also surfaces in $\mathbb{R}^3$. The proofs follow the main outline of the polynomial partitioning proof of Szemerédi-Trotter. All of the results we discuss here were proven earlier using the cutting method, mostly in [**CEGSW**]. There is a lot more information about this type of problem in [**PS**].

EXERCISE 10.1. If Γ is a set of N unit circles in the plane and $\mathcal{S}$ is a set of S points in the plane, prove that

$$|I(\mathcal{S}, \Gamma)| \lesssim N^{2/3} S^{2/3} + N + S.$$

As a corollary, show that a set of N points in the plane determines $\lesssim N^{4/3}$ unit distances.

EXERCISE 10.2. Next suppose that Γ is a set of N circles in the plane and $\mathcal{S}$ is a set of S points in the plane. Check that a set of 3 points lies on a unique circle. Using this observation and a double counting bound, prove that

$$|I(\Gamma, \mathcal{S})| \leq S^3 + N.$$

(There is a slightly different bound which follows from Theorem 8.9: $|I(\Gamma, \mathcal{S})| \leq N^{2/3} S + N$.)

In particular, if $N > S^3$, then $|I(\Gamma, \mathcal{S})| \lesssim N$. This bound is sharp because we can easily arrange that every curve of Γ passes through at least one point of $\mathcal{S}$.

Now combine this counting bound with polynomial partitioning to give an improved estimate. Show that

$$|I(\Gamma, \mathcal{S})| \leq S^{3/5} N^{4/5} + N + S.$$

As a corollary, show that

$$|P_r(\Gamma)| \lesssim N^2 r^{-5/2}.$$

The same arguments apply to parabolas as well as circles. Here a parabola means a curve of the form $y = ax^2 + bx + c$.

Remark: These bounds on the incidence problem for parabolas and circles were first proven in [**CEGSW**] using the cutting method. In the early 2000's, [**ArSh**] proved somewhat better bounds. These bounds don't match any examples. All known bounds give the same estimates for circles and parabolas. On the other hand, the best known examples of parabolas have far more incidences than for circles.

EXERCISE 10.3. In this exercise, we describe a set of parabolas and a set of points with many incidences. Suppose that $\mathcal{S}$ is a grid of points $(x, y) \in \mathbb{Z}^2$ with $|x| \leq X$ and $|y| \leq Y$. Let Γ be the set of all parabolas of the form $y = ax^2 + bx + c$ where $(a, b, c) \in \mathbb{Z}^3$ with $|a| \leq A$, $|b| \leq B$, $|c| \leq C$. By adjusting the parameters X, Y, A, B, C, try to find an example with many incidences.

EXERCISE 10.4. Suppose that Γ is a set of N irreducible algebraic curves in the plane of degree at most d. In other words, each curve in Γ is the zero set of an irreducible polynomial $Q \in \mathrm{Poly}_d(\mathbb{R}^2)$. Using the Bezout theorem, show that for any set of $d^2 + 1$ points in $\mathbb{R}^2$, there is at most one curve $\gamma \in \Gamma$ containing all the points.

If $N \geq S^{d^2+1}$, prove that

$$|I(\Gamma, \mathcal{S})| \lesssim N.$$

Combining this bound and polynomial partitioning, prove the following estimate for the number of incidences: setting $k = d^2 + 1$,

$$|I(\Gamma, \mathcal{S})| \lesssim S^{\frac{k}{2k-1}} N^{\frac{2k-2}{2k-1}} + S + N.$$

Remark: For $d \geq 3$, this bound was improved in [**WYZ**]. For $d = 2$, this is the best known bound at the present time. For all $d \geq 2$, the best current bound does not match any known examples.

Exercise 10.5. In the arguments above, we do polynomial partitioning using polynomials of large degree – the degree is typically a polynomial in terms of the size of Γ and the size of $\mathcal{S}$. The paper [**SoTa**] gave an alternate proof of a slightly weaker estimate which only uses polynomials of a large constant degree. This is a useful technique in some other problems, especially in higher dimensions. We will use this approach later in this chapter to study lines in $\mathbb{R}^3$.

For any $\varepsilon > 0$, constant degree partitioning leads to a proof of the following weak version of Szemerédi-Trotter: if $\mathfrak{L}$ is a set of L lines in the plane, and $\mathcal{S}$ is a set of S points in the plane, then

$$|I(\mathcal{S}, \mathfrak{L})| \leq C(\varepsilon) \left(L^{\frac{2}{3}+\varepsilon} S^{\frac{2}{3}+\varepsilon} + S + L \right).$$

Let $D = D(\varepsilon)$ be a large constant degree. Using polynomial partitioning with polynomials of degree at most D, together with induction on the size of $\mathfrak{L}$ and $\mathcal{S}$, prove this inequality.

Next we turn to incidence problems about surfaces in $\mathbb{R}^3$. We start with 2-planes and then discuss unit 2-spheres and 2-spheres.

In order to apply the partitioning method to 2-planes in $\mathbb{R}^3$, we need to estimate the number of components of $\mathbb{R}^3 \setminus Z(P)$ which a 2-plane Π may enter. The Harnack inequality leads to such an estimate.

Theorem 10.12. (Harnack inequality) If $P \in \mathrm{Poly}_D(\mathbb{R}^2)$, then $\mathbb{R}^2 \setminus Z(P)$ has $\lesssim D^2$ connected components.

Therefore, if $P \in \mathrm{Poly}_D(\mathbb{R}^3)$, and $\Pi \subset \mathbb{R}^3$ is a 2-plane, then $\Pi \setminus Z(P)$ has $\lesssim D^2$ connected components. In particular, a 2-plane Π enters $\lesssim D^2$ connected components of $\mathbb{R}^3 \setminus Z(P)$.

In a later exercise, we will sketch a proof of this theorem. Before that, let us apply it to give an incidence estimate. Suppose that Γ is a set of N 2-planes in $\mathbb{R}^3$ and $\mathcal{S}$ is a set of S points in $\mathbb{R}^3$. We might like to estimate $|I(\Gamma, \mathcal{S})|$. However, this problem does not turn out to be interesting. It may happen that there is a line $l \subset \mathbb{R}^3$ so that every point of $\mathcal{S}$ lies in l, and every plane of Γ contains l. In this case, $|I(\Gamma, \mathcal{S})| = SN$, and this is the maximum possible value. In order to get an interesting question, we need to modify the hypotheses. One possible modification is to bound the number of planes of Γ containing any line.

Exercise 10.6. In this exercise, we prove a slightly weaker version of a theorem from [**EGS**] about the incidences between points and planes. We mentioned this result earlier as Theorem 8.5.

Suppose that Γ is a set of N 2-planes in $\mathbb{R}^3$ where no three 2-planes are collinear. Suppose that $\mathcal{S}$ is a set of S points in $\mathbb{R}^3$.

By a double counting argument, show that

$$|I(\Gamma, \mathcal{S})| \leq 2S^2 + N.$$

Combine this counting bound with polynomial partitioning to prove an incidence estimate: for any $\varepsilon > 0$,

$$I(\Gamma, \mathcal{S}) \leq C(\varepsilon) \left(S^{\frac{4}{5}+\varepsilon} N^{\frac{3}{5}+\varepsilon} + S + N \right).$$

In the proof, use constant degree partitioning as in Exercise 10.5, as well as the Harnack inequality.

(This result was first proven in [**EGS**], using the cutting method. The proof there is a little sharper – avoiding the ε's. This is one of the fairly rare inequalities in higher dimensions which is known to be sharp. The example is described by Apfelbaum and Sharir in Appendix A of [**ApSh**], following some ideas of Brass and Knauer from [**BrKn**].)

Here are some other variations. What would happen if instead of saying that no three two-planes are collinear, we say that there are at most 100 planes of Γ containing any line? We could also replace the condition on planes by a condition on points. What if no three points are collinear? What if at most 100 points lie on any line?)

EXERCISE 10.7. In this exercise, we give a sketch Harnack's inequality, Theorem 10.12. Let P be a non-zero polynomial in $\mathrm{Poly}_D(\mathbb{R}^2)$. We have to estimate the number of connected components of $\mathbb{R}^2 \setminus Z(P)$. The key tool in the estimate is Bezout's theorem.

We will estimate the bounded and unbounded components separately. If S is a circle not contained in $Z(P)$, then $|Z(P) \cap S| \leq 2D$ by the Bezout theorem. Considering large circles, prove that $\mathbb{R}^2 \setminus Z(P)$ has at most $2D$ unbounded components.

Next consider bounded components. A key observation is that inside of each bounded component of $\mathbb{R}^2 \setminus Z(P)$, P must have either a local maximum or a local minimum. So each bounded component must contain a critical point of P, a point where $\partial_1 P = \partial_2 P = 0$. If the polynomials $\partial_1 P$ and $\partial_2 P$ have no common factor, then the Bezout theorem implies that the number of such zeroes is at most $(\mathrm{Deg}\, \partial_1 P)(\mathrm{Deg}\, \partial_2 P) \leq (D-1)^2$. In this case, the number of bounded components of $\mathbb{R}^2 \setminus Z(P)$ is at most $(D-1)^2$.

It may happen that $\partial_1 P$ and $\partial_2 P$ have a common factor. In this case, we consider the critical points of a small perturbation of P. Consider the polynomial $Q = P + w_1 x_1 + w_2 x_2$. If $|w_1|, |w_2|$ are small enough, prove that Q also has a critical point in each connected component of $\mathbb{R}^2 \setminus Z(P)$. If w_1, w_2 are generic, then prove that $\partial_1 Q$ and $\partial_2 Q$ have no common factor.

Similar arguments apply to spheres. For these arguments, we need a version of the Harnack inequality for spheres.

THEOREM 10.13. (Spherical version of the Harnack inequality) If $P \in \mathrm{Poly}_D(\mathbb{R}^3)$, and $S^2 \subset \mathbb{R}^3$ is a sphere (with any center or radius), then $S^2 \setminus Z(P)$ has $\lesssim D^2$ connected components.

The proof of Theorem 10.13 is similar to the proof in the planar case described in the last exercise but technically harder. For a reader interested in the different versions of Bezout's theorem, it would be interesting to work it out.

EXERCISE 10.8. Suppose that Γ is a set of N unit 2-spheres in $\mathbb{R}^3$, and suppose $\mathcal{S}$ is a set of S points in $\mathbb{R}^3$. Check that the intersection of three unit 2-spheres consists of at most two points. In other words, given any three points in $\mathbb{R}^3$, there are at most two spheres $\gamma \in \Gamma$ containing these three points. Using this, show that if $N \geq S^3$, then

$$I(\Gamma, \mathcal{S}) \lesssim N.$$

Combine this counting bound with polynomial partitioning to prove an incidence estimate: for any $\varepsilon > 0$,

$$I(\Gamma, \mathcal{S}) \leq C(\varepsilon) S^{\frac{3}{4}+\varepsilon} N^{\frac{3}{4}+\varepsilon} + S + N.$$

As a corollary, show that N points in $\mathbb{R}^3$ determine at most $N^{\frac{3}{2}+\varepsilon}$ unit distances.

This result was first proven in [**CEGSW**].

EXERCISE 10.9. Now suppose that Γ is a set of N 2-spheres in $\mathbb{R}^3$ (of any radii). Note that the intersection of two 2-spheres is always a circle (or a point). Suppose that $\mathcal{S}$ is a set of S points in $\mathbb{R}^3$ with at most 10 on any circle. Try to estimate $|I(\Gamma, \mathcal{S})|$.

We have seen how polynomial partitioning (or the cutting method) gives interesting estimates about a wide variety of incidence problems. However, in all of the problems we considered in this section, the bounds from polynomial partitioning are not believed to be sharp.

10.6. First estimates for lines in $\mathbb{R}^3$

In the rest of this chapter, we use polynomial partitioning to study lines in $\mathbb{R}^3$. We are now ready to return to the question at the start of the chapter: "Suppose that $\mathfrak{L}$ is a set of lines in $\mathbb{R}^3$ with at most B lines in any algebraic surface of degree at most D. What is the maximum possible size of $P_r(\mathfrak{L})$?"

We will focus in this section on 2-rich points, and we begin with some examples. A set of B lines in a plane can have $\binom{B}{2} \sim B^2$ 2-rich points. A set of B lines in a regulus can also have $\sim B^2$ 2-rich points. If we choose L/B planes or reguli in general position and select B lines from each, we get a set of L lines with $\sim (L/B)B^2 \sim BL$ 2-rich points, and with at most B lines in any low degree surface. For $B \geq L^{1/2}$, we will eventually prove that this bound is sharp. The methods in this book hit a barrier at $L^{3/2}$ 2-rich points, and we won't be able to prove any estimate better than this even if B is very small. We will eventually prove that the number of 2-rich points is $\lesssim BL + L^{3/2}$, and in this chapter, we will prove that the number of 2-rich points is $\lesssim_\varepsilon BL + L^{(3/2)+\varepsilon}$.

We need to find a way to use the hypothesis that not many lines of $\mathfrak{L}$ lie in any algebraic surface of degree at most D. We will use polynomial partitioning with a polynomial $P \in \mathrm{Poly}_D(\mathbb{R}^n)$, and the hypothesis implies that there are few lines in $Z(P)$. We will use this estimate to help control the number of r-rich points in $Z(P)$. And we will use induction to control the number of r-rich points in the cells. With these tools, we can give a rather short proof of an estimate for 2-rich points. This is not the strongest estimate we will prove, but we start here because the proof is short and clear, and later we will add more ideas to improve it.

PROPOSITION 10.14. For any $\varepsilon > 0$, there is a degree $D = D(\varepsilon)$ and a constant $C(\varepsilon)$ so that the following holds. If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with at most B lines in any algebraic surface of degree $\leq D$, then

$$|P_2(\mathfrak{L})| \leq C(\varepsilon) B^{(1/2)-\varepsilon} L^{(3/2)+\varepsilon}.$$

This result is most interesting if B is small. For instance if $B \lesssim \log L$, then $|P_2(\mathfrak{L})| \lesssim_\varepsilon L^{(3/2)+\varepsilon}$, which is nearly the best known bound in this situation. If B is larger, this result is not as good. In connection with the distinct distance

problem, we are interested in $B = L^{1/2}$. In this case Proposition 10.14 gives $|P_2(\mathfrak{L})| \lesssim_\varepsilon L^{(7/4)+\varepsilon}$, but we will eventually prove that $|P_2(\mathfrak{L})| \lesssim L^{3/2}$.

Proof. We will choose the degree $D = D(\varepsilon)$ below. We do polynomial partitioning with degree D for the set of r-rich points $P_2(\mathfrak{L})$. By Theorem 10.3, there exists a non-zero $P \in \mathrm{Poly}_D(\mathbb{R}^3)$ so that $\mathbb{R}^3 \setminus Z(P)$ is a disjoint union of $\sim D^3$ cells O_i, and each cell contains $\lesssim D^{-3}|P_2(\mathfrak{L})|$ points of $P_2(\mathfrak{L})$.

If most of the points of $P_2(\mathfrak{L})$ lie in the union of the cells O_i, we proceed as follows. We let $\mathfrak{L}_i \subset \mathfrak{L}$ denote the set of lines in $\mathfrak{L}$ that intersect O_i. Notice that $P_2(\mathfrak{L}) \cap O_i \subset P_2(\mathfrak{L}_i)$. Since most of the points of $P_2(\mathfrak{L})$ lie in the cells O_i, there must be $\sim D^3$ cells O_i obeying the following inequality:

$$|P_2(\mathfrak{L})| \lesssim D^3 |P_2(\mathfrak{L}_i)|. \tag{10.2}$$

Each line $l \in \mathfrak{L}$ enters at most $D+1$ of the cells. Therefore, $\sum_i |\mathfrak{L}_i| \leq (D+1)L$. Since there are $\sim D^3$ cells O_i obeying equation 10.2, one of them must have $|\mathfrak{L}_i| \lesssim D^{-2}L$. We now study this cell O_i.

We will study $P_2(\mathfrak{L}_i)$ by induction. We will choose D sufficiently large to guarantee that $|\mathfrak{L}_i| < |\mathfrak{L}|$, so we can use induction on the number of lines. By induction we can assume that $|P_2(\mathfrak{L}_i)| \leq C(\varepsilon) B^{(1/2)-\varepsilon} |\mathfrak{L}_i|^{(3/2)+\varepsilon}$. Assembling all the information, we have the following estimate:

$$|P_2(\mathfrak{L})| \lesssim D^3 |P_2(\mathfrak{L}_i)| \lesssim D^3 C(\varepsilon) B^{(1/2)-\varepsilon} (D^{-2}L)^{(3/2)+\varepsilon}.$$

The total power of D in this inequality is $-2\varepsilon < 0$. (The exponent $(3/2) + \varepsilon$ was chosen exactly to make this happen.) We can rewrite the last inequality as:

$$|P_2(\mathfrak{L})| \leq C D^{-2\varepsilon} C(\varepsilon) B^{(1/2)-\varepsilon} L^{(3/2)+\varepsilon}.$$

In this equation C is an absolute constant. We now choose $D(\varepsilon)$ sufficiently large so that $CD^{-2\varepsilon} \leq 1$. As long as the majority of points of $P_2(\mathfrak{L})$ lie in the union of the cells O_i, the induction closes and we get the desired bound for $|P_2(\mathfrak{L})|$.

Suppose on the other hand that at least half the points of $P_2(\mathfrak{L})$ lie on the surface $Z(P)$. In this case, we will estimate $|P_2(\mathfrak{L})|$ directly. We let $\mathfrak{L}_Z$ be the set of lines of $\mathfrak{L}$ that lie in $Z(P)$. By hypothesis, $|\mathfrak{L}_Z| \leq B$. If $x \in P_2(\mathfrak{L}) \cap Z(P)$, then either $x \in P_2(\mathfrak{L}_Z)$, or x lies in a line of $\mathfrak{L} \setminus \mathfrak{L}_Z$. Since $|\mathfrak{L}_Z| \leq B$, $|P_2(\mathfrak{L}_Z)| \leq B^2$. On the other hand, each line $l \in \mathfrak{L} \setminus \mathfrak{L}_Z$ intersects $Z(P)$ in at most D points. The total number of such intersection points is $\leq DL$. In total,

$$|P_2(\mathfrak{L}) \cap Z(P)| \leq B^2 + DL.$$

We can assume $B \leq L$, and if we choose $C(\varepsilon) \geq 4D(\varepsilon)$, then we get $|P_2(\mathfrak{L})| \leq C(\varepsilon) B^{(1/2)-\varepsilon} L^{(3/2)+\varepsilon}$ as desired. □

We discuss the proof a little. There was some algebra in the proof, and it may not be totally clear where the expression $B^{(1/2)-\varepsilon} L^{(3/2)+\varepsilon}$ comes from. Suppose we try to prove that $|P_2(\mathfrak{L})| \lesssim F(B, L)$ for some function F. Because of easy examples, we need to have $F(B, L) \gtrsim B^2 + L$ for any $B \leq L$. In order for the inductive step to work, we need another condition on $F(B, L)$. Ignoring small factors, this condition roughly says that $F(B, L) \geq D^3 F(B, LD^{-2})$. The smallest function that obeys these two conditions is $B^{1/2} L^{3/2}$. We need a little extra room in the argument because of the constant factors, and this leads to the right-hand side $C(\varepsilon) B^{(1/2)-\varepsilon} L^{(3/2)+\varepsilon}$.

10.7. An estimate for r-rich points

Essentially the same proof also leads to an estimate for $|P_r(\mathfrak{L})|$. To estimate the number of r-rich points in $Z(P)$, we use the Szemerédi-Trotter theorem. We state this estimate as a lemma.

LEMMA 10.15. Suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$, $P \in \mathrm{Poly}_D(\mathbb{R}^3)$, and that $Z(P)$ contains at most B lines of $\mathfrak{L}$. Then

$$|P_r(\mathfrak{L}) \cap Z(P)| \lesssim DLr^{-1} + B^2 r^{-3}.$$

PROOF. Let $\mathfrak{L}_Z$ be the set of lines of $\mathfrak{L}$ that lie in $Z(P)$. If $x \in P_r(\mathfrak{L}) \cap Z(P)$, then either x lies in at least $r/10$ lines of $\mathfrak{L} \setminus \mathfrak{L}_Z$, or else x lies in at least $(9/10)r$ lines of $\mathfrak{L}_Z$. We estimate the number of points of each type.

Since each line of $\mathfrak{L} \setminus \mathfrak{L}_Z$ intersects $Z(P)$ in at most D points, the number of points of the first type is at most $DL(r/10)^{-1} = 10DLr^{-1}$.

Since there are at most B lines in $\mathfrak{L}_Z$, the number of points of the second type is $\lesssim B^2 r^{-3} + Br^{-1}$. The term Br^{-1} is dominated by DLr^{-1}, so we don't need to include it on the right-hand side. □

Let $I(B, D, L, r)$ be defined to be the maximum possible size of $|P_r(\mathfrak{L})|$ for a set of L lines with at most B lines in any algebraic surface of degree at most D.

EXERCISE 10.10. Adapting the proof of Proposition 10.14, prove the following inductive estimate for $I(B, D, L, r)$:

$$I(B, D, L, r) \leq C\Big(D^3 I(B, D, CLD^{-2}, r) + DLr^{-1} + B^2 r^{-3}\Big). \tag{10.3}$$

We also know $I(B, D, L, r)$ for $r \geq 2L^{1/2}$. Lemma 7.3 says that if $r \geq 2L^{1/2}$, then $|P_r(\mathfrak{L})| \leq 2Lr^{-1}$. Therefore, if $r \geq 2L^{1/2}$, then $I(B, D, L, r) \leq 2Lr^{-1}$. We restate the result here and give a slightly different proof, which will be a model for another proof a little later in the chapter.

LEMMA 10.16. If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^n$ and $r > 2L^{1/2}$, then $|P_r(\mathfrak{L})| \leq 2Lr^{-1}$. Therefore, if $r \geq 2L^{1/2}$, then $I(B, D, L, r) \leq 2Lr^{-1}$.

PROOF. Let $P_r(\mathfrak{L})$ be $x_1, x_2, .., x_M$, with $M = |P_r(\mathfrak{L})|$. Now x_1 lies in at least r lines of $\mathfrak{L}$. The point x_2 lies in at least $(r-1)$ lines of $\mathfrak{L}$ that did not contain x_1. More generally, the point x_j lies in at least $r - (j-1)$ lines of $\mathfrak{L}$ that did not contain any of the previous points $x_1, ..., x_{j-1}$. Therefore, we have the following inequality for the total number of lines:

$$L \geq \sum_{j=1}^{M} \max(r - j, 0).$$

If $M \geq r/2$, then we would get $L \geq (r/2)(r/2) = r^2/4$. But by hypothesis, $r > 2L^{1/2}$, giving a contradiction. Therefore, $M < r/2$, and we get $L \geq M(r/2)$ which proves the proposition. □

EXERCISE 10.11. Combining the inductive estimate in Equation 10.3 with Lemma 10.16, prove the following theorem.

PROPOSITION 10.17. For any $\varepsilon > 0$, there exists a degree $D = D(\varepsilon)$ and a constant $C(\varepsilon)$ so that the following holds. Suppose that $\mathfrak{L}$ is a set of L lines in

$\mathbb{R}^3$ with at most B lines in any algebraic surface of degree $\leq D$. Then for any $2 \leq r \leq 2L^{1/2}$,

$$|P_r(\mathfrak{L})| \leq C(\varepsilon) B^{(1/2)-\varepsilon} L^{(3/2)+\varepsilon} r^{-2}.$$

If B is very small, then this estimate is close to the best known bound. For instance, if $B \lesssim \log L$, then we get the bound $|P_r(\mathfrak{L})| \lesssim_\varepsilon L^{(3/2)+\varepsilon} r^{-2}$. The best known bound in this situation is $|P_r(\mathfrak{L})| \lesssim L^{3/2} r^{-2}$, as explained in Chapter 12. But for larger B, the bound in Proposition 10.17 is not so sharp.

EXERCISE 10.12. Suppose that $\mathfrak{L}$ is a set of lines in $\mathbb{R}^n$ with at most B lines in any algebraic hypersurface of degree at most D. Try to estimate $|P_r(\mathfrak{L})|$.

10.8. The main theorem

Now we turn to the main theorem of the chapter, Theorem 10.1. We restate the theorem for convenience.

THEOREM. For any $\varepsilon > 0$, there are constants $D(\varepsilon)$ and $C(\varepsilon)$ so that the following holds. If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with at most $L^{(1/2)+\varepsilon}$ lines in any algebraic surface of degree $\leq D(\varepsilon)$, then

$$|P_r(\mathfrak{L})| \leq C(\varepsilon) L^{(3/2)+\varepsilon} r^{-2} + 2Lr^{-1}.$$

If $r > 2L^{1/2}$, then $|P_r(\mathfrak{L})| \leq 2Lr^{-1}$ by the double counting argument in Lemma 10.16, so the interesting case is in the range $2 \leq r \leq 2L^{1/2}$.

We will prove this theorem using induction on L and polynomial partitioning. The proof follows the method from the last two sections, but with a new wrinkle. If we directly follow the method from the last section, we run into the following issue. After polynomial partitioning, we consider the lines $\mathfrak{L}_i$ that intersect a cell O_i. Since $\mathfrak{L}_i \subset \mathfrak{L}$, we know that at most $L^{(1/2)+\varepsilon}$ lines of $\mathfrak{L}_i$ lie in any algebraic surface of degree at most D, but we don't know that at most $|\mathfrak{L}_i|^{(1/2)+\varepsilon}$ lines of $\mathfrak{L}_i$ lie in any algebraic surface of degree at most D. So we cannot directly apply induction to $\mathfrak{L}_i$. We have to somehow deal with low degree algebraic surfaces that contain more than $|\mathfrak{L}_i|^{(1/2)+\varepsilon}$ lines of $\mathfrak{L}_i$. The new wrinkle is a way to deal with these surfaces.

We will actually prove a slightly stronger theorem, because a slightly stronger theorem makes the induction work better. This stronger theorem roughly says that if a set of lines $\mathfrak{L}$ in $\mathbb{R}^3$ has more than $L^{(3/2)+\varepsilon} r^{-2}$ r-rich points, then most of these points must "come from" a small set of low degree algebraic surfaces. Here is a little notation to help state the theorem: if Z is an algebraic surface in $\mathbb{R}^3$, then we define $\mathfrak{L}_Z \subset \mathfrak{L}$ to be the set of lines in $\mathfrak{L}$ that lie in Z.

THEOREM 10.18. For any $\epsilon > 0$, there are $D(\epsilon)$, and $K(\epsilon)$ so that the following holds. For any $r \geq 2$, let $r' = \lceil (9/10)r \rceil$, the least integer which is at least $(9/10)r$.

If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$, and if $2 \leq r \leq 2L^{1/2}$, then there is a set $\mathcal{Z}$ of algebraic surfaces so that

- Each surface $Z \in \mathcal{Z}$ is an irreducible surface of degree at most D.
- Each surface $Z \in \mathcal{Z}$ contains more than $L^{(1/2)+\epsilon}$ lines of $\mathfrak{L}$.
- $|\mathcal{Z}| \leq 2L^{(1/2)-\epsilon}$.
- $|P_r(\mathfrak{L}) \setminus \cup_{Z \in \mathcal{Z}} P_{r'}(\mathfrak{L}_Z)| \leq K L^{(3/2)+\epsilon} r^{-2}$.

Theorem 10.18 immediately implies Theorem 10.1. Suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with at most $L^{(1/2)+\varepsilon}$ lines in any algebraic surface of degree at most $D(\varepsilon)$. As we remarked above, if $r > 2L^{1/2}$, then we know that $|P_r(\mathfrak{L})| \leq 2Lr^{-1}$ by double counting. If $r \leq 2L^{1/2}$, then we apply Theorem 10.18. Since $\mathfrak{L}$ contains at most $L^{(1/2)+\varepsilon}$ lines in any algebraic surface of degree at most D, the set $\mathcal{Z}$ must be empty and so $|P_r(\mathfrak{L})| \leq KL^{(3/2)+\epsilon}r^{-2}$.

Here is an outline of the proof of Theorem 10.18. We use a polynomial partitioning argument to cut $\mathbb{R}^3$ into cells O_i. We use induction to study the lines of $\mathfrak{L}$ that enter each cell. For each cell, we get a set of surfaces $\mathcal{Z}_i$ that accounts for all but a few of the r-rich points in O_i. We let $\tilde{\mathcal{Z}}$ be the union of all the surfaces from the different cells together with the irreducible components of the polynomial partitioning surface. The good news is that the surfaces in $\tilde{\mathcal{Z}}$ account for almost all of the r-rich points of $\mathfrak{L}$: as long as $D(\varepsilon)$ is large enough, a computation similar to the last two sections shows that

$$|P_r(\mathfrak{L}) \setminus \cup_{Z \in \tilde{\mathcal{Z}}} P_{r'}(\mathfrak{L}_Z)| \leq (1/100)KL^{(3/2)+\epsilon}r^{-2}. \tag{10.4}$$

But there is also some bad news. A surface $Z \in \mathcal{Z}_i$ must contain more than $|\mathfrak{L}_i|^{(1/2)+\varepsilon}$ lines of $\mathfrak{L}_i$, but it doesn't have to contain more than $L^{(1/2)+\varepsilon}$ lines of $\mathfrak{L}$. If Z contains $\leq L^{(1/2)+\varepsilon}$ lines of $\mathfrak{L}$, then we are not allowed to include Z in $\mathcal{Z}$. Another piece of bad news is that the number of surfaces in $\tilde{\mathcal{Z}}$ is too large. It turns out that the number of surfaces in $\tilde{\mathcal{Z}}$ could be bigger than $CDL^{(1/2)-\varepsilon}$ for a large constant C. On the other hand, to close the induction, we need to choose a set of surfaces $\mathcal{Z}$ with $|\mathcal{Z}| \leq 2L^{(1/2)-\varepsilon}$. The factor CD may not seem that large, but inductive proofs are very delicate, and $\tilde{\mathcal{Z}}$ contains far more surfaces than we are allowed to put into $\mathcal{Z}$.

At the moment, we might worry that $\tilde{\mathcal{Z}}$ consists of $1000DL^{(1/2)-\varepsilon}$ surfaces, each containing $L^{(1/2)+\varepsilon} - 1$ lines of $\mathfrak{L}$. The key new idea in this section is that this scenario is impossible (for large L). Recall that the surfaces in $\tilde{\mathcal{Z}}$ are irreducible algebraic surfaces of degree $\leq D$. By the Bezout theorem, Theorem 6.7, the intersection of two such surfaces contains at most D^2 lines. But in the scenario above, a simple counting argument shows that two of the surfaces of $\tilde{\mathcal{Z}}$ would have to share more than D^2 lines. More generally, we will prove the following estimate about the number of surfaces that contain many lines of $\mathfrak{L}$.

LEMMA 10.19. Let $\mathfrak{L}$ be a set of L lines in $\mathbb{R}^3$. Suppose Z_j are irreducible algebraic surfaces of degree at most D, each containing at least A lines of $\mathfrak{L}$. If $A > 2DL^{1/2}$, then the number of surfaces Z_j is at most $2L/A$.

PROOF OF LEMMA 10.19. The proof of Lemma 10.19 is a double counting argument, closely analogous to the proof of Lemma 10.16. Suppose that $Z_j = Z(P_j)$ for an irreducible polynomial P_j of degree at most D. Since the surfaces Z_j are distinct, no two polynomials of the polynomials P_j can have a common factor. By the Bezout theorem for lines, Theorem 6.7, the number of lines in $Z(P_{j_1}) \cap Z(P_{j_2})$ is at most D^2 for any $j_1 \neq j_2$.

The surface Z_1 contains at least A lines of $\mathfrak{L}$. The surface Z_2 contains at least $A - D^2$ lines of $\mathfrak{L}$ that are not in Z_1. In general, the surface Z_j contains at least $A - (j-1)D^2$ lines that are not in the previous Z_j. If the number of surfaces Z_j

is J, then we get the following inequality:

$$\sum_{j=1}^{J} \max(A - (j-1)D^2, 0) \leq L. \qquad (*)$$

The rest of the proof is just a computation to see what Inequality $(*)$ tells us about J. If $j \leq (1/2)AD^{-2}$, then $A - jD^2 \geq A/2$. If $J \geq (1/2)AD^{-2}$, we see that $L \geq (1/2)AD^{-2}(A/2) = (1/4)A^2D^{-2}$. Since $A > 2DL^{1/2}$, this inequality gives the contradiction $L > L$. Therefore, $J \leq (1/2)AD^{-2}$. Therefore, all the terms on the left-hand side of $(*)$ are at least $A/2$, and we get $J(A/2) \leq L$. And so $J \leq 2L/A$ as desired. □

We will use Lemma 10.19 to control $\tilde{\mathcal{Z}}$. We let $\mathcal{Z} \subset \tilde{\mathcal{Z}}$ be the subset of surfaces that contain many lines of $\mathfrak{L}$:

$$\mathcal{Z} := \{Z \in \tilde{\mathcal{Z}} | Z \text{ contains more than } L^{(1/2)+\epsilon} \text{ lines of } \mathfrak{L}\}.$$

Just by definition, we know that each surface $Z \in \mathcal{Z}$ contains more than $L^{(1/2)+\varepsilon}$ lines of $\mathfrak{L}$. And Lemma 10.19 gives the desired bound on $|\mathcal{Z}|$:

$$|\mathcal{Z}| \leq 2L^{(1/2)-\epsilon}. \qquad (10.5)$$

To close the induction, it remains to prove that $\mathcal{Z}$ accounts for almost all the r-rich points of $\mathfrak{L}$: we have to bound $|P_r(\mathfrak{L}) \setminus \cup_{Z \in \mathcal{Z}} P_{r'}(\mathfrak{L}_Z)|$. Given Inequality 10.4, we just have to check that the surfaces in $\tilde{\mathcal{Z}}/\mathcal{Z}$ did not contribute too much to controlling the r-rich points of $\mathfrak{L}$. More precisely we will prove that

$$\sum_{Z \in \tilde{\mathcal{Z}} \setminus \mathcal{Z}} |P_{r'}(\mathfrak{L}_Z)| \leq (1/100)KL^{(3/2)+\epsilon}r^{-2}. \qquad (10.6)$$

To prove this estimate, we sort $\tilde{\mathcal{Z}}$ according to the number of lines in each surface. We define

$$\tilde{\mathcal{Z}}_s := \{Z \in \tilde{\mathcal{Z}} \text{ so that } |\mathfrak{L}_Z| \in [2^s, 2^{s+1})\}.$$

A surface $Z \in \tilde{\mathcal{Z}} \setminus \mathcal{Z}$ contains at most $L^{(1/2)+\varepsilon}$ lines, and so it belongs to some $\tilde{\mathcal{Z}}_s$ with $2^s \leq L^{(1/2)+\varepsilon}$. Now for each s, we use Lemma 10.19 to estimate $|\tilde{\mathcal{Z}}_s|$, and for each $Z \in \tilde{\mathcal{Z}}_s$, we use the Szemerédi-Trotter theorem to estimate $|P_{r'}(\mathfrak{L}_Z)|$. Adding up the terms gives Inequality 10.6.

We have now finished our outline of the plan and assembled our tools, and we are ready to prove Theorem 10.18.

Proof. We start with some minor book-keeping to describe the constants D and K and to explain the base of the induction. We remark that if $\epsilon \geq 1/2$ then the theorem is trivial: we can take $\mathcal{Z}$ to be empty, and it is easy to check that $|P_r(\mathfrak{L})| \leq 2L^2r^{-2}$. (This follows from Szemerédi-Trotter, which gives a stronger estimate. But it also follows from a simple double-counting argument.) So we can assume that $\epsilon \leq 1/2$.

We will choose D a large constant depending on ϵ and then we will choose K a large constant depending on ϵ and D. As long as these are large enough at certain points in the proof, the argument goes through.

The proof is by induction on L. By choosing K large, we can assume that the theorem holds when L is small. As long as $K \geq 4L^2$, Theorem 10.18 is trivial: we take $\mathcal{Z}$ to be empty, and we observe that $|P_r(\mathfrak{L})| \leq L^2$. Since $r \leq 2L^{1/2}$,

$L^2 \leq KL^{(3/2)+\varepsilon} r^{-2}$. We choose $K \geq 4(2D)^{2/\varepsilon}$, so that the Theorem holds whenever $L \leq (2D)^{1/\varepsilon}$. This is the base of the induction.

Now we turn to the induction, which is the heart of the matter.

10.8.1. Building $\tilde{\mathcal{Z}}$. Let S be any subset of $P_r(\mathfrak{L})$. An important case is $S = P_r(\mathfrak{L})$, but we will have to consider other sets as well. We use Theorem 10.3 to do a polynomial partitioning of the set S with a polynomial of degree at most D. The polynomial partitioning theorem, Theorem 10.3, says that there is a non-zero polynomial P of degree at most D so that

- $\mathbb{R}^3 \setminus Z(P)$ is the union of at most CD^3 disjoint open cells O_i, and
- for each cell O_i, $|S \cap O_i| \leq CD^{-3}|S|$.

We define $\mathfrak{L}_i \subset \mathfrak{L}$ to be the set of lines from $\mathfrak{L}$ that intersect the open cell O_i. We note that $S \cap O_i \subset P_r(\mathfrak{L}_i)$. If a line does not lie in $Z(P)$, then it can have at most D intersection points with $Z(P)$, which means that it can enter at most $D+1$ cells O_i. So each line of $\mathfrak{L}$ intersects at most $D+1$ cells O_i. Therefore, we get the following inequality:

$$\sum_i |\mathfrak{L}_i| \leq (D+1)L \leq 2DL. \tag{10.7}$$

Let $\beta > 0$ be a large parameter that we will choose below. We say that a cell O_i is β-good if

$$|\mathfrak{L}_i| \leq \beta D^{-2} L. \tag{10.8}$$

The number of β-bad cells is at most $2\beta^{-1}D^3$. Each cell contains at most $CD^{-3}|S|$ points of S. Therefore, the bad cells all together contain at most $C\beta^{-1}|S|$ points of S. We now choose β so that $C\beta^{-1} \leq (1/100)$. β is an absolute constant, independent of ϵ. We now have the following estimate:

(10.9) The union of the bad cells contains at most $(1/100)|S|$ points of S.

For each good cell O_i, we apply induction to understand $\mathfrak{L}_i$. By choosing D sufficiently large, we can guarantee that for each good cell, $|\mathfrak{L}_i| \leq (1/2)L$. Now there are two cases, depending on whether $r \leq 2|\mathfrak{L}_i|^{1/2}$.

If $r \leq 2|\mathfrak{L}_i|^{1/2}$, then we can apply the inductive hypothesis. In this case, we see that there is a set $\mathcal{Z}_i$ of irreducible algebraic surfaces of degree at most D with the following two properties:

$$|\mathcal{Z}_i| \leq 2|\mathfrak{L}_i|^{(1/2)-\epsilon} \leq 2(\beta D^{-2}L)^{(1/2)-\epsilon}. \tag{10.10}$$

$$|P_r(\mathfrak{L}_i) \setminus \cup_{Z \in \mathcal{Z}_i} P_{r'}(\mathfrak{L}_Z)| \leq K|\mathfrak{L}_i|^{(3/2)+\epsilon} r^{-2} \leq K(\beta D^{-2}L)^{(3/2)+\epsilon} r^{-2}.$$

Because $S \cap O_i \subset P_r(\mathfrak{L}_i)$, we see that

$$|(S \cap O_i) \setminus \cup_{Z \in \mathcal{Z}_i} P_{r'}(\mathfrak{L}_Z)| \leq C_1 K D^{-3-2\epsilon} L^{(3/2)+\epsilon} r^{-2}. \tag{10.11}$$

On the other hand, if $r > 2|\mathfrak{L}_i|^{1/2}$, then we define $\mathcal{Z}_i$ to be empty, and Lemma 10.16 gives the bound

$$|S \cap O_i| \leq |P_r(\mathfrak{L}_i)| \leq 2|\mathfrak{L}_i| r^{-1} \leq 2Lr^{-1} \leq 4L^{3/2} r^{-2}. \tag{10.12}$$

By choosing K sufficiently large compared to D, we can arrange that $4L^{3/2}r^{-2}$ $\leq C_1 K D^{-3-2\epsilon} L^{(3/2)+\epsilon} r^{-2}$. Therefore, inequality 10.11 holds for the good cells with

$r > 2|\mathfrak{L}_i|^{1/2}$ as well as the good cells with $r \leq 2|\mathfrak{L}_i|^{1/2}$. We sum this inequality over all the good cells:

$$\sum_{O_i \text{ good}} |(S \cap O_i) \setminus \cup_{Z \in \mathcal{Z}_i} P_{r'}(\mathfrak{L}_Z)| \leq CD^3 \cdot C_1 K D^{-3-2\epsilon} L^{(3/2)+\epsilon} r^{-2}$$
$$\leq C_2 D^{-2\epsilon} K L^{(3/2)+\epsilon} r^{-2}.$$

We choose $D(\epsilon)$ large enough so that $C_2 D^{-2\epsilon} \leq (1/400)$. Therefore, we get the following:

$$\sum_{O_i \text{ good}} |(S \cap O_i) \setminus \cup_{Z \in \mathcal{Z}_i} P_{r'}(\mathfrak{L}_Z)| \leq (1/400) K L^{(3/2)+\epsilon} r^{-2}. \tag{10.13}$$

We have studied the points of S in the good cells. Next we study the points of S in the zero set of the partioning polynomial $Z(P)$. Let Z_j be an irreducible component of $Z(P)$. If $x \in S \cap Z_j$, but $x \notin P_{r'}(\mathfrak{L}_{Z_j})$, then x must lie in at least $r/10$ lines of $\mathfrak{L} \setminus \mathfrak{L}_{Z_j}$. Each line of $\mathfrak{L}$ that is not contained in Z_j has at most $\mathrm{Deg}(Z_j)$ intersection points with Z_j. Therefore,

$$|(S \cap Z_j) \setminus P_{r'}(\mathfrak{L}_{Z_j})| \leq 10 r^{-1} (\mathrm{Deg}\, Z_j) L.$$

If $\{Z_j\}$ are all the irreducible components of $Z(P)$, then we see that

$$|(S \cap Z(P)) \setminus \cup_j P_{r'}(\mathfrak{L}_{Z_j})| \leq 10 r^{-1} D L.$$

We choose $K = K(\epsilon, D)$ sufficiently large so that $10D \leq (1/800)K$. Since $r \leq 2L^{1/2}$, we have

$$|(S \cap Z(P)) \setminus \cup_j P_{r'}(\mathfrak{L}_{Z_j})| \leq (1/800) K L r^{-1} \leq (1/400) K L^{3/2} r^{-2}. \tag{10.14}$$

Now we define $\tilde{\mathcal{Z}}_S$ to be the union of $\mathcal{Z}_i$ over all the good cells O_i together with all the irreducible components Z_j of $Z(P)$. Each surface in $\tilde{\mathcal{Z}}_S$ is an algebraic surface of degree at most D. By equation 10.10, we have the following estimate for $|\tilde{\mathcal{Z}}_S|$:

$$|\tilde{\mathcal{Z}}_S| \leq CD^3 (\beta D^{-2} L)^{(1/2)-\epsilon} + D \leq CD^3 L^{(1/2)-\epsilon}. \tag{10.15}$$

(We could have written something a little smaller than the right-hand side, with a more complicated power of D depending on ε, but notice that our bound is at least $CD^2 L^{(1/2)-\epsilon}$.)

Summing the contribution of the bad cells in equation 10.9, the contribution of the good cells in equation 10.13, and the contribution of the cell walls in equation 10.14, we get:

$$|S \setminus \cup_{Z \in \tilde{\mathcal{Z}}_S} P_{r'}(\mathfrak{L}_Z)| \leq (1/100)|S| + (1/200) K L^{(3/2)+\epsilon} r^{-2}. \tag{10.16}$$

If we didn't have the $(1/100)|S|$ term coming from the bad cells, we could simply take $S = P_r(\mathfrak{L})$ and $\tilde{\mathcal{Z}} = \tilde{\mathcal{Z}}_S$. Because of the bad cells, we need to run the above construction repeatedly. This is a minor detail which we didn't mention in the outline above.

Let $S_1 = P_r(\mathfrak{L})$, and let $\tilde{\mathcal{Z}}_{S_1}$ be the set of surfaces constructed above. Now we define $S_2 = S_1 \setminus \cup_{Z \in \tilde{\mathcal{Z}}_{S_1}} P_{r'}(\mathfrak{L}_Z)$. We iterate this procedure, defining

$$S_{j+1} := S_j \setminus \cup_{Z \in \tilde{\mathcal{Z}}_{S_j}} P_{r'}(\mathfrak{L}_Z).$$

Each set S_j is a subset of $P_r(\mathfrak{L})$. By Equation 10.15, each set of surfaces $\tilde{\mathcal{Z}}_{S_j}$ has cardinality at most $CD^3L^{(1/2)-\epsilon}$. Iterating equation 10.16 we see:

$$|S_{j+1}| \leq (1/100)|S_j| + (1/200)KL^{(3/2)+\epsilon}r^{-2}. \tag{10.17}$$

We define $J = 1000 \log L$. Since $|S_1| = |P_r(\mathfrak{L})| \leq L^2$, the iterative formula in equation 10.17 implies that

$$|S_J| \leq (1/100)KL^{(3/2)+\epsilon}r^{-2}. \tag{10.18}$$

We define $\tilde{\mathcal{Z}} = \cup_{j=1}^{J-1} \tilde{\mathcal{Z}}_{S_j}$. This set of surfaces has the following properties. Since each set $\tilde{\mathcal{Z}}_{S_j}$ has at most $CD^3L^{(1/2)-\epsilon}$ surfaces, we get:

$$|\tilde{\mathcal{Z}}| \lesssim D^3L^{(1/2)-\epsilon} \log L. \tag{10.19}$$

Our bound for $|\tilde{\mathcal{Z}}|$ is $\log L$ times bigger than we could have gotten if there were no bad cells. It turns out that this factor of $\log L$ does not have much effect on the estimates later in the proof.

Also, $P_r(\mathfrak{L}) \setminus \cup_{Z \in \tilde{\mathcal{Z}}} P_{r'}(\mathfrak{L}_Z) = S_J$, and so equation 10.18 gives:

$$|P_r(\mathfrak{L}) \setminus \cup_{Z \in \tilde{\mathcal{Z}}} P_{r'}(\mathfrak{L}_Z)| \leq (1/100)KL^{(3/2)+\epsilon}r^{-2}. \tag{10.20}$$

This finishes our construction of $\tilde{\mathcal{Z}}$. Next we prune $\tilde{\mathcal{Z}}$ down to our desired set of surfaces $\mathcal{Z}$.

10.8.2. Pruning $\tilde{\mathcal{Z}}$. We define

$$\mathcal{Z} := \{Z \in \tilde{\mathcal{Z}} | Z \text{ contains at least } L^{(1/2)+\epsilon} \text{ lines of } \mathfrak{L}\}.$$

To close our induction, we have to check two properties of $\mathcal{Z}$.

(1) $|\mathcal{Z}| \leq 2L^{(1/2)-\epsilon}$.
(2) $|P_r(\mathfrak{L}) \setminus \cup_{Z \in \mathcal{Z}} P_{r'}(\mathfrak{L}_Z)| \leq KL^{(3/2)+\epsilon}r^{-2}$.

To prove item (1), we apply Lemma 10.19 to the set of surfaces $\mathcal{Z}$ with $A = L^{(1/2)+\epsilon}$. To apply the lemma, we need to know that $A > 2DL$, which is equivalent to $L^{\varepsilon} > 2D$. We can assume that $L^{\epsilon} > 2D$, because the case of $L^{\epsilon} \leq 2D$ was the base of our induction, and we handled it by choosing K sufficiently large. Therefore, the hypotheses of Lemma 10.19 are satisfied. The lemma tells us that $|\mathcal{Z}| \leq 2L^{(1/2)-\epsilon}$.

Now we turn to item (2). We proved above that the surfaces of $\tilde{\mathcal{Z}}$ account for all but a small number of r-rich points. We made this precise in equation 10.20:

$$|P_r(\mathfrak{L}) \setminus \cup_{Z \in \tilde{\mathcal{Z}}} P_{r'}(\mathfrak{L}_Z)| \leq (1/100)KL^{(3/2)+\epsilon}r^{-2}.$$

Therefore, it suffices to check that

$$\sum_{Z \in \tilde{\mathcal{Z}} \setminus \mathcal{Z}} |P_{r'}(\mathfrak{L}_Z)| \leq (1/100)KL^{(3/2)+\epsilon}r^{-2}. \tag{10.21}$$

We sort $\tilde{\mathcal{Z}} \setminus \mathcal{Z}$ according to the number of lines in each surface. For each integer $s \geq 0$, we define:

$$\tilde{\mathcal{Z}}_s := \{Z \in \tilde{\mathcal{Z}} \text{ so that } |\mathfrak{L}_Z| \in [2^s, 2^{s+1})\}.$$

Since each surface of $\tilde{\mathcal{Z}}$ with at least $L^{(1/2)+\epsilon}$ lines of $\mathfrak{L}$ lies in $\mathcal{Z}$, we see that:

$$\tilde{\mathcal{Z}} \setminus \mathcal{Z} \subset \bigcup_{2^s \leq L^{(1/2)+\epsilon}} \tilde{\mathcal{Z}}_s. \tag{10.22}$$

Now we can break up $\sum_{Z\in\tilde{\mathcal{Z}}\setminus\mathcal{Z}}|P_{r'}(\mathfrak{L}_Z)|$ into contributions from different values of s:

$$\sum_{Z\in\tilde{\mathcal{Z}}\setminus\mathcal{Z}}|P_{r'}(\mathfrak{L}_Z)|\leq\sum_{2^s\leq L^{(1/2)+\epsilon}}\left(\sum_{Z\in\tilde{\mathcal{Z}}_s}|P_{r'}(\mathfrak{L}_Z)|\right). \tag{10.23}$$

For each $Z\in\tilde{\mathcal{Z}}_s$, we use the Szemerédi-Trotter theorem, Theorem 7.1, to bound $P_{r'}(\mathfrak{L}_Z)$. Since $Z\in\tilde{\mathcal{Z}}_s$, $|\mathfrak{L}_Z|\leq 2^{s+1}$. Since $r'\geq(9/10)r$, Szemerédi-Trotter gives the bound $|P_{r'}(\mathfrak{L}_Z)|\leq C(2^{2s}r^{-3}+2^sr^{-1})$. Plugging this estimate into Inequality 10.23, we get

$$\sum_{Z\in\tilde{\mathcal{Z}}\setminus\mathcal{Z}}|P_{r'}(\mathfrak{L}_Z)|\leq C\sum_{2^s\leq L^{(1/2)+\epsilon}}|\tilde{\mathcal{Z}}_s|\left(2^{2s}r^{-3}+2^sr^{-1}\right). \tag{10.24}$$

Next, we estimate $|\tilde{\mathcal{Z}}_s|$. If $2^s>2DL^{1/2}$, then Lemma 10.19 gives the estimate $|\tilde{\mathcal{Z}}_s|\leq 2L2^{-s}$. This allows us to estimate the contribution to 10.24 from s in the range $2DL^{1/2}<2^s\leq L^{(1/2)+\epsilon}$ as follows:

$$\sum_{2DL^{1/2}<2^s\leq L^{(1/2)+\epsilon}}|\tilde{\mathcal{Z}}_s|\left(2^{2s}r^{-3}+2^sr^{-1}\right)\leq\sum_{2^s\leq L^{(1/2)+\epsilon}}(2L2^{-s})\left(2^{2s}r^{-3}+2^sr^{-1}\right)\leq$$

$$\leq C\sum_{2^s\leq L^{(1/2)+\epsilon}}(L2^sr^{-3}+Lr^{-1})\leq C(L^{(3/2)+\epsilon}r^{-3}+L(\log L)r^{-1}).$$

Since $r\leq 2L^{1/2}$, this last expression is $\leq CL^{(3/2)+\epsilon}r^{-2}$. In summary, we see that

$$\sum_{2DL^{1/2}<2^s\leq L^{(1/2)+\epsilon}}|\tilde{\mathcal{Z}}_s|\left(2^{2s}r^{-3}+2^sr^{-1}\right)\leq CL^{(3/2)+\epsilon}r^{-2}. \tag{10.25}$$

Next we consider the contribution to 10.24 from s in the range $2^s\leq 2DL^{1/2}$. For s in this range, we use Equation 10.19 to estimate $|\tilde{\mathcal{Z}}_s|\leq|\tilde{\mathcal{Z}}|\lesssim D^3L^{(1/2)-\varepsilon}\log L$.

$$\sum_{2^s\leq 2DL^{1/2}}|\tilde{\mathcal{Z}}_s|\left(2^{2s}r^{-3}+2^sr^{-1}\right)\lesssim D^3\left(L^{(1/2)-\epsilon}\log L\right)\left(Lr^{-3}+L^{1/2}r^{-1}\right).$$

Since $r\leq 2L^{1/2}$, this is $\lesssim D^3L^{3/2}r^{-2}$. In summary, we see that

$$\sum_{2^s\leq 2DL^{1/2}}|\tilde{\mathcal{Z}}_s|\left(2^{2s}r^{-3}+2^sr^{-1}\right)\lesssim D^3L^{3/2}r^{-2}. \tag{10.26}$$

Combining Inequalities 10.24, 10.25, and 10.26, we get the bound

$$\sum_{Z\in\tilde{\mathcal{Z}}\setminus\mathcal{Z}}|P_{r'}(\mathfrak{L}_Z)|\leq D^3L^{(3/2)+\epsilon}r^{-2}.$$

If we choose $K=K(\epsilon,D)$ sufficiently large, then

$$\sum_{Z\in\tilde{\mathcal{Z}}\setminus\mathcal{Z}}|P_{r'}(\mathfrak{L}_Z)|\leq(1/100)KL^{(3/2)+\epsilon}r^{-2}.$$

This is Inequality 10.21. As we discussed above, it implies item (2):

$$|P_r(\mathfrak{L})\setminus\cup_{Z\in\mathcal{Z}}P_{r'}(\mathfrak{L}_Z)|\leq KL^{(3/2)+\epsilon}r^{-2}.$$

This estimate closes the induction, finishing the proof of Theorem 10.18. □

https://doi.org/10.1090//ulect/064/11

CHAPTER 11

Combinatorial structure, algebraic structure, and geometric structure

The next three chapters are about the structure of configurations of lines with many r-rich points. If $\mathfrak{L}$ is a set of lines in $\mathbb{R}^3$ and P is a non-zero polynomial of minimal degree that vanishes on $\mathfrak{L}$, we will see that the combinatorics of $\mathfrak{L}$, the algebraic properties of P, and the geometry of $Z(P)$ are all connected with each other.

For example, suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with $\gg L^{3/2}$ 3-rich points. The only examples that we have seen occur when the lines of $\mathfrak{L}$ cluster into a small number of planes. Using polynomial partitioning, we proved that if $|P_3(\mathfrak{L})| \geq C(\varepsilon)L^{(3/2)+\varepsilon}$, then the lines of $\mathfrak{L}$ cluster into algebraic surfaces of degree at most $D(\varepsilon)$. In this chapter, we will prove the following sharper theorem:

THEOREM 11.1. There is a constant K so that the following holds. If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with $|P_3(\mathfrak{L})| \geq KL^{3/2}$, then there is a plane that contains at least $10L^{1/2}$ lines of $\mathfrak{L}$.

The proof of this theorem uses completely different methods from the polynomial partitioning arguments in the last chapter. We suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with at least $KL^{3/2}$ 3-rich points. This hypothesis describes the combinatorial structure of $\mathfrak{L}$. We prove that there is a polynomial P vanishing on $\mathfrak{L}$ with a surprisingly small degree. This step shows that $\mathfrak{L}$ has a special algebraic structure. With the help of this algebraic structure, we study the geometry of $Z(P)$. We prove that $Z(P)$ has many flat points, and eventually that $Z(P)$ contains a plane which contains many of the lines of $\mathfrak{L}$. In summary, combinatorial structure leads to algebraic structure which leads to geometric structure.

In Chapter 12, combining these techniques with polynomial partitioning, we will prove a sharper theorem about r-rich points for all $r \geq 3$.

11.1. Structure for configurations of lines with many 3-rich points

Theorem 11.1 gives a lot of information about the structure of configurations of lines in $\mathbb{R}^3$ with many 3-rich points. At first sight, the conclusion may look a little weak. We assumed that $\mathfrak{L}$ has more than $L^{3/2}$ 3-rich points, and the conclusion tells us that there is a plane with at least $L^{1/2}$ lines of $\mathfrak{L}$. If a plane contains $L^{1/2}$ lines of $\mathfrak{L}$ then the lines in the plane can only have at most L 3-rich points, so it seems that they only account for a small fraction of all the 3-rich points of $\mathfrak{L}$.

To see that exponent of $L^{1/2}$ in the conclusion is sharp, we consider the following example. Consider L/A planes in general position. Suppose that $\mathfrak{L}$ consists of A lines from each plane. Within each plane, we can arrange these A lines in a grid pattern so that they contribute $\sim A^2$ 3-rich points. Then the total number of 3-rich

points is $\sim LA$. If we take $A = CL^{1/2}$ for a large constant C, then $\mathfrak{L}$ will have $KL^{3/2}$ 3-rich points, and each plane will contain at most $CL^{1/2} \lesssim L^{1/2}$ lines of $\mathfrak{L}$.

In this example, each plane contains at most $\sim L^{1/2}$ lines of $\mathfrak{L}$, but there are many such planes, and all together they account for almost all of the 3-rich points of $\mathfrak{L}$. Theorem 11.1 implies that every set of lines with many 3-rich points has this type of structure.

COROLLARY 11.2. If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$, then there is a set of planes $\Pi_1, ..., \Pi_S$ with $S \leq L^{1/2}$, and there are disjoint subsets $\mathfrak{L}_i \subset \mathfrak{L}$ so that the lines of $\mathfrak{L}_i$ are contained in Π_i so that

$$|P_3(\mathfrak{L}) \setminus \cup_i P_3(\mathfrak{L}_i)| \leq KL^{3/2}.$$

The corollary shows that if $|P_3(\mathfrak{L})|$ is much larger than $KL^{3/2}$ then almost all of the 3-rich points are accounted for by subsets of lines concentrated in various 2-dimensional planes.

PROOF. We prove the corollary by induction on L. If $|P_3(\mathfrak{L})| \leq KL^{3/2}$, then the estimate holds immediately (with no subsets $\mathfrak{L}_i$). If $|P_3(\mathfrak{L})| > KL^{3/2}$, then Theorem 11.1 implies that there is some plane Π_1 containing at least $10L^{1/2}$ lines of $\mathfrak{L}$. Define $\mathfrak{L}_1$ to be the set of lines of $\mathfrak{L}$ that lie in Π_1.

If $|P_3(\mathfrak{L}) \setminus P_3(\mathfrak{L}_i)| \leq KL^{3/2}$, then we are done. Otherwise, we let $\mathfrak{L}' := \mathfrak{L} \setminus \mathfrak{L}_1$. We know that

$$|\mathfrak{L}'| \leq L - 10L^{1/2}.$$

By induction on L, we know that the corollary is true for $\mathfrak{L}'$. We note that $|\mathfrak{L}'|^{1/2} < L^{1/2} - 1$. Applying the corollary to $\mathfrak{L}'$, we get planes $\Pi_2, ..., \Pi_S$ with $S \leq L^{1/2}$ and disjoint subsets $\mathfrak{L}_i \subset \mathfrak{L}'$ for $2 \leq i \leq s$, so that the lines of $\mathfrak{L}_i$ lie in Π_i and so that

$$\left|P_3(\mathfrak{L}') \setminus \cup_{i=2}^{S} P_3(\mathfrak{L}_i)\right| \leq K|\mathfrak{L}'|^{3/2} \leq K(L - 10L^{1/2})^{3/2}. \tag{11.1}$$

We have now defined the planes $\Pi_1, ..., \Pi_S$ and the sets $\mathfrak{L}_1, ..., \mathfrak{L}_S$. Since $\mathfrak{L}_1$ and $\mathfrak{L}'$ are disjoint, it follows that $\mathfrak{L}_1$ is disjoint from any of the sets $\mathfrak{L}_2, ..., \mathfrak{L}_S$, and so the sets $\mathfrak{L}_i$ are disjoint as claimed. We just have to estimate $|P_3(\mathfrak{L}) \setminus \cup_i P_3(\mathfrak{L}_i)|$.

Suppose that $x \in P_3(\mathfrak{L})$. Either $x \in P_3(\mathfrak{L}_1)$ or $x \in P_3(\mathfrak{L}')$ or x lies in at least one line of $\mathfrak{L}_1$ and of $\mathfrak{L}'$. In this last case, x must lie in the plane Π_1. Every line of $\mathfrak{L}'$ intersects Π_1 in at most one point, and so the number of 3-rich points in this last case is at most L. In other words:

$$|P_3(\mathfrak{L}) \setminus (P_3(\mathfrak{L}_1) \cup P_3(\mathfrak{L}'))| \leq L. \tag{11.2}$$

Combining Equations 11.1 and 11.2, we get

$$|P_3(\mathfrak{L}) \setminus \cup_i P_3(\mathfrak{L}_i)| \leq K(L - 10L^{1/2})^{3/2} + L \leq KL^{3/2}.$$

This closes the induction and finishes the proof. □

For configurations of lines in $\mathbb{R}^3$ with significantly more than $L^{3/2}$ 3-rich points, Corollary 11.2 gives a detailed description of the structure. In Chapter 13, we will prove a similar structure theorem for configurations of lines with significantly more than $L^{3/2}$ 2-rich points. This structure theorem will involve planes and degree 2 algebraic surfaces.

It would be interesting to try to go below $KL^{3/2}$. Can we prove a structure theorem for configurations of lines with $(1/10)L^{3/2}$ 3-rich points? How about $L^{1.49}$ 3-rich points? In Section 7.5 we discussed overdetermined and underdetermined

problems in incidence geometry. Asking a set of L lines in $\mathbb{R}^3$ to have $L^{1.49}$ 3-rich points is heavily overdetermined in this sense, and it is plausible that there may be some structure theroem for such configurations, but all the methods we know so far break down in this regime. We will come back to this question at the end of Chapter 13, after we have seen all of the incidence geometry techniques in the book.

Now we turn to the proof of Theorem 11.1.

11.2. Algebraic structure and degree reduction

We begin by thinking about the algebraic complexity of a set and the idea of algebraic structure. Let $\mathbb{F}$ be a field and let $S \subset \mathbb{F}^n$. We define $\mathrm{Deg}(S)$ to be the minimal degree of a non-zero polynomial that vanishes on S. At the beginning of the book, in Proposition 2.1, we used parameter counting to prove an estimate for $\mathrm{Deg}(S)$:

PROPOSITION 11.3. Suppose that $S \subset \mathbb{F}^n$, and that $|S| < \mathrm{Dim}\,\mathrm{Poly}_D(\mathbb{F}^n) = \binom{D+n}{n}$. Then $\mathrm{Deg}(S) \leq D$.

This result is sharp. We will check in the appendix that if $M \geq \binom{D+n}{n}$, then a generic set S of M points in $\mathbb{F}^n$ has degree $> D$. A generic set S has degree $\sim |S|^{1/n}$. If $\mathrm{Deg}(S)$ is much smaller than $|S|^{1/n}$, then we say that the set S has algebraic structure.

We can define the degree of a set of lines in a similar way. Suppose that $\mathfrak{L}$ is a set of lines in $\mathbb{F}^n$. We define $\mathrm{Deg}(\mathfrak{L})$ to be the minimal degree of a non-zero polynomial that vanishes on each line of $\mathfrak{L}$. Any set of L lines in $\mathbb{F}^n$ has degree $\lesssim L^{\frac{1}{n-1}}$.

PROPOSITION 11.4. If $\mathfrak{L}$ is a set of L lines in $\mathbb{F}^n$, and if $(D+1)L < \mathrm{Dim}\,\mathrm{Poly}_D(\mathbb{F})^n = \binom{D+n}{n}$, then $\mathrm{Deg}(\mathfrak{L}) \leq D$. In particular, $\mathrm{Deg}(\mathfrak{L}) \leq (2n+1)L^{\frac{1}{n-1}}$.

PROOF. Let S be a set of points with $D+1$ points on each line of $\mathfrak{L}$, and with $|S| \leq (D+1)|\mathfrak{L}|$. By Proposition 11.3, there is a non-zero $P \in \mathrm{Poly}_D(\mathbb{F}^n)$ which vanishes on S. By the vanishing lemma, it vanishes on each line of $\mathfrak{L}$.

Now we check that $\mathrm{Deg}(\mathfrak{L}) \leq (2n+1)L^{\frac{1}{n-1}}$. This part is just a computation. We take D to be the greatest integer less than $(2n+1)L^{\frac{1}{n-1}}$, and so we have $D \geq (2n)L^{\frac{1}{n-1}}$. It suffices to check that $(D+1)L \leq D^n n^{-n} < \binom{D+n}{n}$.

We have

$$(D+1)L \leq (2n+1)L^{\frac{n}{n-1}} \leq (2n+1)^n n^{-n} L^{\frac{n}{n-1}} = D^n n^{-n}.$$

□

If the degree is below the bound given by Proposition 11.4, then the set of lines $\mathfrak{L}$ has algebraic structure. In particular, if $n = 3$, then any set of L lines has degree $\leq 7L^{1/2}$, and if the degree is significantly less than $L^{1/2}$, then the lines have algebraic structure.

Now we come to a connection between combinatorial structure and algebraic structure. We show that a set of lines in $\mathbb{F}^3$ with many intersection points has small degree.

PROPOSITION 11.5. Let $\mathfrak{L}$ be a set of lines in $\mathbb{F}^3$. Suppose that each line of $\mathfrak{L}$ contains at least A points of $P_2(\mathfrak{L})$. Then $\mathrm{Deg}(\mathfrak{L}) \lesssim L/A$.

Proposition 11.5 is an important philosophical point in the polynomial method. The hypothesis that each line of $\mathfrak{L}$ has $\geq A$ intersection points with other lines of $\mathfrak{L}$ describes the combinatorics of $\mathfrak{L}$. By Proposition 11.5, this combinatorial structure implies that $\mathfrak{L}$ has algebraic structure: it implies that there is a polynomial of surprisingly low degree that vanishes on $\mathfrak{L}$. Once we know about this polynomial, it is reasonable to try to exploit this algebaric structure to study $\mathfrak{L}$.

We consider a couple examples to put Proposition 11.5 in context. If $A = L^{1/2}$, then Proposition 11.5 says that $\mathrm{Deg}(\mathfrak{L}) \leq CL^{1/2}$ for a large constant C. By Proposition 11.4, we already know that $\mathrm{Deg}(\mathfrak{L}) \leq 7L^{1/2}$ for any set of L lines. In this case, the bound from Proposition 11.5 is not interesting. But if A is much larger than $L^{1/2}$, Proposition 11.5 will show that $\mathrm{Deg}(\mathfrak{L})$ is much smaller than $L^{1/2}$, showing that $\mathfrak{L}$ has algebraic structure. The larger A is, the more structure $\mathfrak{L}$ has.

When $A \geq L^{1/2}$, there is a simple example that shows we cannot hope to reduce the degree below $L/(A+1)$. Suppose that $A+1$ divides L. Choose $L/(A+1)$ planes, and let $\mathfrak{L}$ contain $A+1$ lines in each of the planes. Within each plane, we choose the $A+1$ lines in general position, so that each line contains A intersection points with other lines. Let Q_j be the degree 1 polynomial that vanishes on the j^{th} plane. The product $\prod Q_j$ vanishes on $\mathfrak{L}$ and has degree $L/(A+1)$. On the other hand, suppose that P vanishes on $\mathfrak{L}$. For each j, there are $A+1$ lines in $Z(P, Q_j)$. By Bezout's theorem for lines, Theorem 6.7, either $\mathrm{Deg}(P) \geq A+1$ or else Q_j and P have a common factor. Since Q_j is degree 1, the only possible common factor is Q_j. Therefore, either $\mathrm{Deg}\, P \geq A+1$ or else each Q_j divides P. So $\mathrm{Deg}\, P \geq \min(A+1, L/(A+1))$. If $A \geq L^{1/2}$, then $\mathrm{Deg}\, P \geq L/(A+1)$. Therefore, Proposition 11.5 is sharp up to a constant factor for all $A \geq L^{1/2}$.

We prove Proposition 11.5 in the next section. The proof uses ideas from the finite field Nikodym proof. I call this argument the contagious vanishing argument.

11.3. The contagious vanishing argument

Here is the main idea of the proof of Proposition 11.5. Let $\mathfrak{L}$ be a set of L lines in $\mathbb{F}^3$ where each line contains $\geq A$ intersection points with other lines. By parameter counting, we can find a polynomial P of degree $D \sim L/A$ that vanishes on $\sim D^2$ lines of $\mathfrak{L}$. In the interesting cases, D^2 will be much smaller than L, so our polynomial only vanishes on a small fraction of the lines. But the vanishing of P is 'contagious': if a line l has $> D$ intersection points with lines where P vanishes, then P vanishes on l also. 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. In our case, we are assuming that each line has $\geq A$ intersection points with other lines. For an appropriate choice of D, the vanishing of P starts on D^2 lines and then spreads to all of the lines.

Next let's do a simple heuristic calculation to figure out how small we can expect to make D. Initially, P vanishes on $\sim D^2$ lines of $\mathfrak{L}$. Let's suppose that we choose these D^2 lines randomly. Let's imagine that the vanishing set of P is colored red, so that we have $\sim D^2$ red lines. Now consider a line $l \in \mathfrak{L}$, and let's estimate the expected number of intersection points between l and these D^2 red lines. For each intersection point between l and $\mathfrak{L}$, the probability that P vanishes at the intersection point is $\gtrsim D^2/L$. The number of intersection points along l is A. So the expected number of red intersection points along l is $\gtrsim AD^2L^{-1}$. If this

expected number is $> 10D$, then we can expect P to vanish on most lines of $\mathfrak{L}$. This condition is $AD^2L^{-1} > CD$ for a universal constant C. Doing a little algebra, it suffices to take any D obeying $D > CL/A$. Therefore, we heuristically expect to do degree reduction with a degree $D \lesssim L/A$. We will see below that this is correct.

In the heuristic above, we discussed the expected value of various quantities. In the full proof, we will need to know that these quantities are close to their expected values with high probability. In particular, we will use the following lemma.

LEMMA 11.6. (Probability lemma) Let S be a set of N elements. Let $X \subset S$ be a random subset where each element of S is included in X independently with probability p. The expected size of X is pN.

(1) $\mathbb{P}[|X| > 2pN] \leq \exp(-\frac{1}{100}pN)$.
(2) $\mathbb{P}[|X| < (1/2)pN] \leq \exp(-\frac{1}{100}pN)$.

The lemma says that the size of $|X|$ is close to the expected value pN almost all the time. It is an example of a large deviation bound. See Appendix A of [**AlSp**] for a good introduction to this type of estimate. We will also give a self-contained proof of Lemma 11.6 at the end of the section.

Now we can begin the formal proof of Proposition 11.5.

PROOF. Let $D \geq 1000$ be a degree which we will choose later. Let p be the number $(1/20)D^2/L$. We form a subset $\mathfrak{L}_0 \subset \mathfrak{L}$ by including each line independently with probability p. The expected number of lines in $\mathfrak{L}_0$ is $pL = (1/20)D^2$. With high probability, the size of $\mathfrak{L}_0$ is at most $(1/10)D^2$. More precisely, because $D \geq 1000$, Lemma 11.6 implies that $|\mathfrak{L}_0| \leq (1/10)D^2$ with probability at least $\frac{99}{100}$. As long as $|\mathfrak{L}_0| \leq (1/10)D^2$, we can find a non-zero polynomial P of degree $\leq D$ that vanishes on the lines of $\mathfrak{L}_0$.

Fix a line $l \in \mathfrak{L}$. It contains $\geq A$ intersection points with other lines of $\mathfrak{L}$. Each of these intersection points has a probability $\geq p$ of lying in a line of $\mathfrak{L}_0 \setminus \{l\}$. These events are independent. The expected number of points of l lying in lines of $\mathfrak{L}_0$ is $E \geq Ap = (1/20)D^2A/L$.

We now choose D in the range $(10^6 - 1)L/A \leq D \leq 10^6 L/A$. An easy calculation shows that $E \geq 10^4 D$.

If l intersects $\mathfrak{L}_0$ in $\geq D + 1$ points, then $P = 0$ on l. But by the probability lemma, Lemma 11.6, the probability that l intersects $\mathfrak{L}_0$ in $\leq D$ points is $\leq exp(-\frac{1}{100}E) \leq \exp(-100D) \leq \exp(-10^7 L/A)$.

If $L/A > 10^{-5} \log L$ then the probability that l contains $\leq D$ intersection points with $\mathfrak{L}_0$ is $< L^{-10}$. In this case, with high probability, P vanishes on every line of $\mathfrak{L}$, and we are done. This is the main case.

If $L/A \leq 10^{-5} \log L$, then the probability that l contains $\leq D$ intersection points with $\mathfrak{L}_0$ is $< \exp(-10^7)$. In this case, with high probability, P vanishes on at least $\frac{99}{100}L$ lines of $\mathfrak{L}$. This does not give the conclusion, but it is a good step.

In order to fully handle the case that $L/A \leq 10^{-5} \log L$, we organize our proof by induction on L. We will prove by induction on L that $\mathfrak{L}$ lies in the zero set of P with $\operatorname{Deg} P \leq 10^7 L/A$. The case $L \leq 10^7$ is now trivial, because any 10^7 lines lie in a union of 10^7 planes. This is the base of our induction. Also, if $L/A > 10^{-5} \log L$, then we proved above that $\mathfrak{L}$ lies in the zero set of a polynomial P with $\operatorname{Deg} P \leq 10^6 L/A$.

If $L/A \leq 10^{-5} \log L$, we proved that there is a polynomial P_1 with $\operatorname{Deg} P_1 \leq 10^6 L/A$ so that P_1 vanishes on $\mathfrak{L}_1 \subset \mathfrak{L}$ with $|\mathfrak{L}_1| \geq (99/100)L$. Let $\mathfrak{L}_2 \subset \mathfrak{L}$

be the set lines of $\mathfrak{L}$ on which P_1 does not vanish. We have $|\mathfrak{L}_2| \leq (1/100)L$. Each line of $\mathfrak{L}_2$ has $\leq \operatorname{Deg} P_1$ intersection points with lines of $\mathfrak{L}_1$. But it has $\geq A$ intersection points with lines of $\mathfrak{L}$. Therefore, each line of $\mathfrak{L}_2$ has at least $A - \operatorname{Deg} P_1$ intersection points with other lines of $\mathfrak{L}_2$. We note that $A \geq 100L(\log L)^{-1}$ and $\operatorname{Deg} P_1 \leq 10^6 L/A \leq 10 \log L$, and so $A - \operatorname{Deg} P_1 \geq (9/10)A$. By induction, we see that $\mathfrak{L}_2$ lies in $Z(P_2)$ for a polynomial P_2 with

$$\operatorname{Deg} P_2 \leq 10^7 |\mathfrak{L}_2| (A - \operatorname{Deg} P_1)^{-1} \leq 10^7 (\frac{1}{100} L)(\frac{9}{10} A)^{-1} \leq 10^6 L/A.$$

Now $P = P_1 P_2$ vanishes on L and has degree at most $(10^6 L/A) + (10^6 L/A) \leq 10^7 L/A$, closing the induction. □

To finish the section, we recall and prove the probability lemma that we used above, Lemma 11.6.

LEMMA. (Probability lemma) Let S be a set of N elements. Let $X \subset S$ be a random subset where each element of S is included in X independently with probability p. The expected size of X is pN.

(1) $\mathbb{P}[|X| > 2pN] \leq \exp(-\frac{1}{100} pN)$.
(2) $\mathbb{P}[|X| < (1/2)pN] \leq \exp(-\frac{1}{100} pN)$.

PROOF. We let a_j be 1 if the j^{th} element of S is included in X and 0 otherwise. The functions a_j are independent, and the probability that $a_j = 1$ is p. Also $|X| = \sum_j a_j$.

If f_j are independent functions, then $\mathbb{E}(\prod_j f_j) = \prod_j (\mathbb{E} f_j)$. To prove the lemma, we will apply this equality for well-chosen functions f_j. Since the functions a_j are independent, we could take $f_j = a_j$, giving the equation $\mathbb{E}(\prod_j a_j) = \prod_j (\mathbb{E} a_j)$. But we are trying to study $|X| = \sum_j a_j$, and having information about $\prod_j a_j$ is not obviously helpful. Instead, we will choose $f_j = e^{\beta a_j}$ for some number $\beta \in \mathbb{R}$. Since the functions a_j are independent, the functions $f_j = e^{\beta a_j}$ are also independent. This is a useful choice for f_j because $\prod_j e^{\beta a_j} = e^{\sum_j \beta a_j} = e^{\beta |X|}$. Therefore, we see that for any $\beta \in \mathbb{R}$,

$$\mathbb{E}\left(e^{\beta |X|}\right) = \mathbb{E}\left(\prod_j e^{\beta a_j}\right) = \prod_j \mathbb{E}\left(e^{\beta a_j}\right) = (pe^{\beta} + 1 - p)^N. \tag{11.3}$$

These inequalities give a lot of information about the distribution of $|X|$. In the rest of the proof, we just extract the information from these bounds to control the probability that $|X|$ lies in a certain range.

We would like to bound $\mathbb{P}[|X| > 2pN]$. We can relate this probability to $\mathbb{E}\left(e^{\beta |X|}\right)$ by observing that

$$\mathbb{P}[|X| > 2pN] \cdot e^{2\beta pN} \leq \mathbb{E}\left(e^{\beta |X|}\right).$$

Combining our last two inequalities, we get the following upper bound for the probability that $|X|$ is $> 2pN$:

$$\mathbb{P}[|X| > 2pN] \leq \left[\frac{pe^{\beta} + 1 - p}{e^{2\beta p}}\right]^N.$$

This bound holds for any $\beta \in \mathbb{R}$. To get the best possible bound on the probability that $|X| > 2pN$, we want to choose β to make the fraction in brackets

as small as possible. In particular, we want the fraction in brackets to be less than 1. Taking $\beta = 1$ gives a reasonable estimate: if $\beta = 1$, the fraction in brackets is

$$\frac{1 + p(e-1)}{e^{2p}} \leq \frac{1 + p(e-1)}{1 + 2p} \leq \exp(-p/100).$$

Therefore, inequality 1 holds.

The proof of inequality 2 is similar but we have to choose β differently. First we observe that

$$\mathbb{P}\left[|X| < (1/2)pN\right] e^{(1/2)\beta pN} \leq \mathbb{E}\left(e^{\beta|X|}\right).$$

Combining this observation with Equation 11.3, we get the following upper bound for the probability that $|X|$ is $< (1/2)pN$:

$$\mathbb{P}\left[|X| < (1/2)pN\right] \leq \left[\frac{pe^{\beta} + 1 - p}{e^{(1/2)\beta p}}\right]^N.$$

This bound again holds for any β. We will see that if β is negative and close to zero, then the expression in brackets is less than 1. In particular, we will check that if $\beta = -1/10$, then the expressions is brackets is less than $e^{-p/100}$, and this will prove inequality 2.

We begin by proving some simple estimates for $e^{-\alpha}$ using Taylor's theorem. Let $g(\alpha) = e^{-\alpha}$. We note that $g''(\alpha) \geq 0$ for all α. Therefore, by Taylor's theorem, for all α, $g(\alpha) \geq g(0) + g'(0)\alpha = 1 - \alpha$. Also, $g'''(\alpha) \leq 0$ for all α. Therefore, by Taylor's theorem, for all $\alpha \geq 0$, $g(\alpha) \leq g(0) + g'(0)\alpha + (1/2)g''(\alpha)\alpha^2 = 1 - \alpha + (1/2)\alpha^2$. In summary, for any $\alpha \geq 0$,

$$1 - \alpha \leq e^{-\alpha} \leq 1 - \alpha + (1/2)\alpha^2.$$

If $\beta = -1/10$, then the expression in brackets is

$$\frac{pe^{\beta} + 1 - p}{e^{(1/2)\beta p}} = \frac{pe^{-1/10} + 1 - p}{e^{-(1/20)p}} \leq \frac{1 - (1/10)p + (1/200)p}{1 - (1/20)p} \leq \exp(-p/100).$$

Therefore, inequality 2 holds. □

11.4. Planar clustering

We will now study configurations of lines in $\mathbb{R}^3$ with many triple points. If there are too many triple points, then we will prove that the lines must cluster in planes. Here is our main result.

THEOREM 11.7. (Planar clustering theorem) (similar to results in [**GK1**], [**EKS**]) There is a constant K so that the following holds. Let $\mathfrak{L}$ be a set of L lines in $\mathbb{R}^3$ so that each line contains $\geq A = KL^{1/2}$ points of $P_3(\mathfrak{L})$. Then $\mathfrak{L}$ lies in $\leq KL/A$ planes.

For example, consider $(1/10)L/A$ planes in general position. Suppose that $\mathfrak{L}$ consists of $1000A$ lines from each plane. Within each plane, we can arrange these $10A$ lines in a grid pattern so that each line contains at least A triple points. In coordinates, we can consider the horizontal lines $y = b$ for $b = 1, ..., 2A$, the vertical lines $x = a$ for $a = 1, ..., 2A$, and the diagonal lines $x - y = c$ for $c = -A, ..., A$. An integer point (x, y) is a triple point for this configuration of lines if $1 \leq x \leq 2A$, $1 \leq y \leq 2A$ and $|x - y| \leq A$. Each line in the collection contains at least A triple points.

COROLLARY 11.8. Suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ that contains at most B lines in any plane. If $B \geq L^{1/2}$, then

$$|P_3(\mathfrak{L})| \lesssim BL.$$

PROOF. Let K be the constant from Theorem 11.7. Using induction on L, we will prove that

$$|P_3(\mathfrak{L})| \leq KBL.$$

If $|P_3(\mathfrak{L})| \leq KBL$, there is nothing to prove, so we may assume that $|P_3(\mathfrak{L})| > KBL \geq KL^{3/2}$.

We apply Theorem 11.7 with $A = |P_3(\mathfrak{L})|L^{-1} \geq KL^{1/2}$. If each line of $\mathfrak{L}$ contains $\geq A$ points of $P_3(\mathfrak{L})$, then Theorem 11.7 implies that $\mathfrak{L}$ is contained in at most KL/A planes. Therefore, one plane contains at least A/K lines of $\mathfrak{L}$, and so $A/K \leq B$. In this case, we can bound $|P_3(\mathfrak{L})|$ as follows:

$$|P_3(\mathfrak{L})| = AL = K(A/K)L \leq KBL.$$

On the other hand, suppose that there is a line $l \in \mathfrak{L}$ that contains at most A points of $P_3(\mathfrak{L})$. We let $\mathfrak{L}' := \mathfrak{L} \setminus \{l\}$. Now we bound $|P_3(\mathfrak{L})|$ by induction:

$$|P_3(\mathfrak{L})| \leq A + |P_3(\mathfrak{L}')| \leq |P_3(\mathfrak{L})|L^{-1} + KB(L-1).$$

Rearranging we get $\frac{L-1}{L}|P_3(\mathfrak{L})| \leq KB(L-1)$, and so

$$|P_3(\mathfrak{L})| \leq KBL.$$

□

11.5. Outline of the proof of planar clustering

Suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$, and each line contains $\geq A \geq KL^{1/2}$ triple intersection points. The first key step is degree reduction, Proposition 11.5. This proposition says that $\mathfrak{L}$ lies in $Z(P)$ where P is a polynomial of degree $\lesssim L/A$. Since A is much larger than $L^{1/2}$, $\mathrm{Deg}(\mathfrak{L})$ is lower than the degree of a generic set of L lines. We can assume that P is a minimal degree polynomial that vanishes on $\mathfrak{L}$. Our goal is to prove that P is a product of linear factors. Since $\mathrm{Deg}\, P \lesssim L/A$, it would then follow that $Z(P)$ is a union of $\lesssim L/A$ planes.

The first observation is that for each point $x \in P_3(\mathfrak{L})$, the three lines of $\mathfrak{L}$ thru x influence the local geometry of $Z(P)$. Recall that a point x in a smooth surface in $\mathbb{R}^3$ is called a flat point if there is a plane that is tangent to the surface at x to second order. We will explain this definition in more detail in the next section. In the proof of the joints theorem, we saw that if x is a joint of $\mathfrak{L}$, then it must be a critical point of the polynomial P. If x is not a critical point of P, then the three lines must lie in the tangent space of $Z(P)$. The lines then force $Z(P)$ to be flat at the point x.

LEMMA 11.9. Every point of $P_3(\mathfrak{L})$ is either a critical point or a flat point of $Z(P)$.

The next observation is that being critical and/or flat is contagious. Once $Z(P)$ has many flat points, then the flatness starts to infect other points, and we will eventually prove that every point of $Z(P)$ is flat. This will show that $Z(P)$ is a union of planes.

The crucial reason that being critical and/or flat is contagious is that these geometric conditions are equivalent to an algebraic condition. In algebraic geometry, one tries to find an algebraic way to describe geometric conditions (and vice versa). In particular, we will work out an algebraic way to describe what it means for a point to be flat, and this leads to the following lemma:

LEMMA 11.10. For any polynomial $P \in \text{Poly}_D(\mathbb{R}^3)$, there is a list of nine polynomials, called $SP_1, ..., SP_9$, of degree at most $3D$, so that a point $x \in Z(P)$ is critical or flat if and only if $SP_1(x) = ... = SP_9(x) = 0$.

Now we can see why being critical/flat is contagious. A line $l \in \mathfrak{L}$ contains $\geq A > 3D$ points of $P_3(\mathfrak{L})$. Each of these points is a critical or flat point of $Z(P)$. So at each of these points, each SP_j vanishes. Now by the vanishing lemma, each SP_j vanishes on l. But then every point of l is either critical or flat. With more work, we will eventually see that all the points of $Z(P)$ are either critical or flat.

Once we know that every point of $Z(P)$ is either critical or flat, we will show that $Z(P)$ is a union of planes. If every point of $Z(P)$ is flat, then we can check with a little differential geometry that $Z(P)$ is a union of planes. There could also be critical points in $Z(P)$, and it takes a little extra work to deal with them, but we will still be able to show that $Z(P)$ is a union of planes. The number of planes is at most $\text{Deg}\, P \lesssim L/A$.

This finishes our outline. Now we start to study flat points more carefully so that we will be able to fill in the details.

11.6. Flat points

We study flat points of a smooth submanifold $M^2 \subset \mathbb{R}^3$. We recall two definitions of a flat point. Consider a point $x \in M^2$. After translating and rotating, we can assume that $x = 0$ and that the tangent space of M at x is $x_3 = 0$. In this case, the manifold M is locally described by a graph $x_3 = h(x_1, x_2)$, where $h(0) = 0$ and $\nabla h(0) = 0$. Now the point x is flat if and only if the second derivatives of h vanish at 0.

There is an alternate definition of flat using the normal vector to the manifold. Suppose that $N : M \to S^2$ is the unit normal vector of M. We say that x is flat if the derivative of N vanishes at x. In other words, x is flat if the map $dN_x : T_xM \to T_{N(x)}S^2$ is zero.

We recall that the derivative dN_x can be defined as follows. We write the normal vector in coordinates as $N = (N_1, N_2, N_3)$. If $v = (v_1, v_2, v_3) \in T_xM$, then

$$dN_x(v) := \sum_{i=1}^{3} v_i \partial_i (N_1, N_2, N_3).$$

For more background about smooth submanifolds and derivatives, the reader can consult the first chapter of [**GP**].

We now check that these two definitions are equivalent. As in the first definition, we choose coordinates so that the point x is 0 and so that near x, M is described as a graph $x_3 = h(x_1, x_2)$ with $h(0) = 0$ and $\nabla h(0) = 0$. Next we describe the normal vector to M in terms of the function h. At the point $(x_1, x_2, h(x_1, x_2))$, the vectors $(1, 0, \partial_1 h)$ and $(0, 1, \partial_2 h)$ are tangent to M, and they span the tangent space to M. Therefore, the vector $(-\partial_1 h, -\partial_2 h, 1)$ is normal to M. We can define

the unit normal vector N by normalizing this vector:

$$N(x_1, x_2, h(x_1, x_2)) = \frac{1}{\sqrt{1 + |\nabla h|^2}}(-\partial_1 h, -\partial_2 h, 1).$$

Next we want to compute the derivative $dN_x(v)$ at the point 0, for vectors $v \in T_x M$. At the point $x = 0$, the tangent space of M is the plane $x_3 = 0$, and it is spanned by the vectors $v_1 = (1, 0, 0)$ and $v_2 = (0, 1, 0)$. By the definition of a derivative, we have at the point $x = 0$,

$$dN_x(v_1) := \partial_1 \left(\frac{1}{\sqrt{1 + |\nabla h|^2}}(-\partial_1 h, -\partial_2 h, 1) \right) \Big|_{x_1 = x_2 = 0}.$$

This formula looks complicated, but at the point x=0, we also know that $\partial_1 h(0) = \partial_2 h(0) = 0$. So when we apply the chain rule, almost all of the terms vanish, leaving the following simple formulas for $dN_x(v)$:

$$dN_x(v_1) = (-\partial_1^2 h, -\partial_1 \partial_2 h, 0).$$
$$dN_x(v_2) = (-\partial_1 \partial_2 h, -\partial_2^2 h, 0).$$

From these formulas we see that $dN_x = 0$ if and only if the second derivatives of h vanish at 0.

We have now reviewed everything that we need to know about flat points.

Next we consider polynomials P on $\mathbb{R}^3$. If x is a regular point of $Z(P)$, then $Z(P)$ is a submanifold in a neighborhood of x.

The first connection between combinatorics of lines and the geometry of algebraic surfaces is the following.

LEMMA 11.11. Suppose that x lies in three lines that lie in $Z(P)$. Then x is either a critical point or a flat point of $Z(P)$.

We saw in the proof of the joints theorem that if x lies in three non-coplanar lines in $Z(P)$, then x is a critical point of $Z(P)$. This result refines that earlier result by describing what happens when the three lines are coplanar.

PROOF. Suppose that x is a non-critical point of $Z(P)$. We see that $\nabla P(x)$ vanishes in the direction of each of the three lines, and so all the lines lie in $T_x Z(P)$.

We use the first definition of flatness above. We rotate and translate so that x is at the origin and $Z(P)$ is described by the equation $x_3 = h(x_1, x_2)$, where the tangent plane of $Z(P)$ is given by $x_3 = 0$. We have three lines contained in $Z(P)$ and in the plane $x_3 = 0$. So h vanishes on these three lines. We expand h in a Taylor series, and look at the second-order terms: $h = h_2 + O(|x|^3)$, where $h_2(x_1, x_2)$ is a homogenous polynomial of degree 2. It follows that h_2 vanishes on three lines in the $x_1 x_2$-plane. But h_2 is a degree 2 polynomial, and it then follows from the vanishing lemma that h_2 is identically zero. This means that x is a flat point. □

Being flat is a geometric condition, but there is an essentially equivalent algebraic condition. A basic theme of algebraic geometry is that geometric features of $Z(P)$ are connected with algebraic features of P and vice versa. The algebraic description of flat points is a small example of this type of correspondence.

Suppose that x is a non-critical point of $Z(P)$. Let N be the unit normal to $Z(P)$. In terms of P, the unit normal vector is given by $N = \frac{\nabla P}{|\nabla P|}$, which is well-defined at every regular point. Our second definition says that x is flat iff

$\partial_v N(x) = 0$ for all $v \in T_x Z(P)$. We would like to say that x is flat if and only if certain polynomials vanish at x. We can adapt the definition of a flat point using a couple of tricks.

The components of ∇P are polynomials, but the components of N are not. Therefore, we would like to rewrite the definition of a flat point in terms of ∇P without mentioning N. The first trick is to see that $\partial_v N = 0$ if and only if $\partial_v \nabla P$ is parallel to ∇P, which happens if and only if $\partial_v \nabla P \times \nabla P = 0$. Here $\times$ denotes the cross-product of vectors in $\mathbb{R}^3$.

The second trick is to note that $\{e_j \times \nabla P\}_{j=1,2,3}$ is a spanning set for $T_x Z(P)$. For two vectors, $v, w \in \mathbb{R}^3$, the cross product $v \times w$ is perpencidular to both v and w. In particular, $e_j \times \nabla P(x)$ is perpendicular to $\nabla P(x)$ and so lies in $T_x Z(P)$. It remains to check that the span of $\{e_j \times \nabla P(x)\}_{j=1,2,3}$ contains all of $T_x Z(P)$. Let v and w be an orthonormal basis of $T_x Z(P)$. Of course, v and w lie in the span of $\{e_j\}_{j=1,2,3}$. Therefore, $v \times \nabla P(x)$ and $w \times \nabla P(x)$ lie in the span of $\{e_j \times \nabla P(x)\}_{j=1,2,3}$. Since $v \times \nabla P(x)$ is perpendicular to v and to $\nabla P(x)$, $v \times \nabla P(x)$ must be a non-zero multiple of w. Similarly, $w \times \nabla P(x)$ must be a non-zero multiple of v. Therefore, the span of $\{e_j \times \nabla P(x)\}_{j=1,2,3}$ contains v and w, and so it contains $T_x Z(P)$.

We are now ready to define the polynomials SP:

$$SP(x) = \{(\partial_{e_j \times \nabla P} \nabla P(x)) \times \nabla P(x)\}_{j=1,2,3}.$$

Note that $SP(x)$ is a list of 3 vectors in $\mathbb{R}^3$, so we can think of it as a list of 9 polynomials. Each polynomial has degree $\leq 3 \operatorname{Deg} P$. We write $SP(x) = 0$ if all 9 polynomials vanish at x.

PROPOSITION 11.12. If $x \in Z(P)$ then $SP(x) = 0$ iff x is critical or flat.

PROOF. If x is critical, then $\nabla P(x) = 0$, and so $SP(x) = 0$.

If x is regular and x is a flat point of $Z(P)$, then $\nabla_v N = 0$ for any $v \in T_x Z(P)$. In particular, $\nabla_{e_j \times \nabla P} N$ vanishes at x, for $j = 1, 2, 3$. Now $\nabla P = |\nabla P| N$, and so $\nabla_{e_j \times \nabla P} \nabla P(x)$ is parallel to N and to ∇P. Therefore, $(\nabla_{e_j \times \nabla P} \nabla P) \times \nabla P$ vanishes at x. In other words, $SP(x) = 0$.

On the other hand, suppose that $SP(x) = 0$ and x is not critical. We have to show that x is flat. Since $SP(x) = 0$, we know that $\nabla_{e_j \times \nabla P} \nabla P(x)$ is parallel to $\nabla P(x)$. Since $N = |\nabla P|^{-1} |\nabla P|$, we see that $\nabla_{e_j \times \nabla P} N(x)$ is parallel to $N(x)$. But since $N \cdot N = 1$ globally, it follows that $\nabla_{e_j \times \nabla P} N(x)$ is perpendicular to $N(x)$. It follows that $\nabla_{e_j \times \nabla P} N(x) = 0$ for $j = 1, 2, 3$. But the vectors $e_j \times \nabla P(x)$ span $T_x Z(P)$. So $\nabla_v N(x) = 0$ for all $v \in T_x Z(P)$, and x is flat. □

Next we investigate what happens if SP vanishes at every point of $Z(P)$. Does this mean that $Z(P)$ is a union of planes? This is not literally true because of some degenerate cases. For example, we could have $P = x_1^2 + x_2^2$. In this case $Z(P) \subset \mathbb{R}^3$ is the line $x_1 = x_2 = 0$. In this case, every point of $Z(P)$ is critical. If $Z(P)$ contains a regular point, then we have the following lemma.

LEMMA 11.13. If P is an irreducible polynomial in $\operatorname{Poly}(\mathbb{R}^3)$, and SP vanishes on $Z(P)$, and $Z(P)$ contains a regular point, then P is a degree 1 polynomial and $Z(P)$ is a plane.

PROOF. Let x be a regular point in $Z(P)$. In a neighborhood of x, $Z(P)$ is a flat submanifold. The normal vector is constant, and so the tangent space is

constant, and so this neighborhood is an open subset of a plane. By the vanishing lemma, $Z(P)$ contains the whole plane. Let P_1 be the degree 1 polynomial that vanishes on the plane.

Next we claim that P_1 divides P. After rotation and translation, we can assume that P_1 is the polynoimal x_3. Now we write P in the form

$$P(x_1, x_2, x_3) = x_3 Q(x_1, x_2, x_3) + R(x_1, x_2).$$

By assumption P vanishes on the plane $x_3 = 0$, and so R vanishes at every point $(x_1, x_2) \in \mathbb{R}^2$. By the Schwarz-Zippel lemma (Exercise 2.3), R is the zero polynomial. (This argument is also very similar to the proof of Lemma 2.10.) Since R is zero, we see that $P = x_3 Q$, and so P_1 divides P.

But since P is irreducible, $P = P_1$. □

We can summarize what we learned here in the following:

PLANE DETECTION LEMMA. For any polynomial P in $\mathbb{R}[x_1, x_2, x_3]$, we can associate a list of polynomials SP with the following properties.

(1) If $x \in Z(P)$ then $SP(x) = 0$ iff x is critical or flat.
(2) If x is contained in three lines in $Z(P)$, then $SP(x) = 0$.
(3) $\operatorname{Deg} SP \leq 3 \operatorname{Deg} P$.
(4) If P is irreducible and SP vanishes on $Z(P)$ and $Z(P)$ contains a regular point, then $Z(P)$ is a plane.

11.7. The proof of the planar clustering theorem

Now we are ready to give the proof of Theorem 11.7.

Let K be a sufficiently large constant.

Let $\mathfrak{L}$ be a set of L lines in $\mathbb{R}^3$. Suppose that each line of $\mathfrak{L}$ contains $\geq A \geq KL^{1/2}$ points of $P_3(\mathfrak{L})$. We let P be the minimal degree non-zero polynomial that vanishes on $\mathfrak{L}$. By the degree reduction argument (Proposition 11.5), we know that $\operatorname{Deg} P \lesssim L/A$. Choosing K large enough, $\operatorname{Deg} P \leq 10^{-2} L^{1/2}$.

We factor P into irreducible factors. We have $P = \prod_j P_j$, where P_j is irreducible. We will decompose $\mathfrak{L}$ into subsets $\mathfrak{L}_j$ corresponding to the P_j. First we define $\mathfrak{L}_{mult}$ to be the set of lines that lie in $Z(P_j)$ for multiple j. By the Bezout theorem, Theorem 6.7, $|\mathfrak{L}_{mult}| \leq \sum_{j,j'} \operatorname{Deg} P_j \operatorname{Deg} P_{j'} = (\operatorname{Deg} P)^2 \leq 10^{-4} L$. So most of the lines lie in exactly one $Z(P_j)$.

We define $\mathfrak{L}_j \subset \mathfrak{L}$ to be the set of lines that lie in $Z(P_j)$ and don't lie in any other $Z(P_{j'})$. Since P has minimal degree, each $\mathfrak{L}_j$ is non-empty.

LEMMA 11.14. Each line in $\mathfrak{L}_j$ contains $\geq (99/100)A$ points of $P_3(\mathfrak{L}_j)$.

PROOF. Let $l \in \mathfrak{L}_j$. By definition of $\mathfrak{L}_j$, for any $j' \neq j$, $P_{j'}$ does not vanish everywhere on l. So $P_{j'}$ vanishes at $\leq \operatorname{Deg} P_{j'}$ points of l. Therefore, there are $\leq \operatorname{Deg} P$ points of l where $P_{j'}$ vanishes for some $j' \neq j$. But $\deg P \leq (1/100)A$. So l contains $\geq (99/100)A$ points of $P_3(\mathfrak{L})$ that don't lie in any other $Z(P_{j'})$. We claim that each of these points lies in $P_3(\mathfrak{L}_j)$. Let x be a point of $l \cap P_3(\mathfrak{L})$, with $P_{j'}(x) \neq 0$ for all $j' \neq j$. The point x lies in at least two other lines of $\mathfrak{L}$, l_1 and l_2. These lines lie in $Z(P)$, but they don't lie in $Z(P_{j'})$ for any $j' \neq j$. Therefore, they lie in $Z(P_j)$, and so they belong to $\mathfrak{L}_j$. □

Since P is the minimal degree polynomial that vanishes on $\mathfrak{L}$, it follows that P_j is the minimal degree polynomial that vanishes on $\mathfrak{L}_j$. By Proposition 11.5, $\operatorname{Deg} P_j \lesssim |\mathfrak{L}_j|/A$. Choosing K large enough, this implies that

$$\operatorname{Deg} P_j \leq 10^{-2}|\mathfrak{L}_j|^{1/2}.$$

At each point $x \in P_3(\mathfrak{L}_j)$, $SP_j(x) = 0$. Now we will use the idea of contagious structure to show that SP_j vanishes on $Z(P_j)$. For each line $l \in \mathfrak{L}_j$, l contains $\geq (99/100)A > 3 \operatorname{Deg} P_j$ points of $P_3(\mathfrak{L}_j)$. We know that $SP_j = 0$ at each of these points, and $\operatorname{Deg} SP_j \leq 3 \operatorname{Deg} P_j$, and so SP_j vanishes on each $l \in \mathfrak{L}_j$.

At this point we use that P_j is irreducible. By the Bezout theorem for lines, Theorem 6.7, either SP_j vanishes on $Z(P_j)$ or else $Z(SP_j) \cap Z(P_j)$ contains $\leq (\operatorname{Deg} SP_j)(\operatorname{Deg} P_j)$ lines. But $(\operatorname{Deg} SP_j)(\operatorname{Deg} P_j) \leq 3(\operatorname{Deg} P_j)^2 < |\mathfrak{L}_j|$, and SP_j and P_j both vanish on every line of $\mathfrak{L}_j$. So we conclude that SP_j vanishes on $Z(P_j)$.

Next we want to show that $Z(P_j)$ contains a regular point. If each point of $P_3(\mathfrak{L}_j)$ were a critical point of P_j, then we would see that ∇P_j vanished on all the lines of $\mathfrak{L}_j$. But $\operatorname{Deg} \partial_i P_j < \operatorname{Deg} P_j$. Since P_j is a minimal degree (non-zero) polynomial that vanishes on $\mathfrak{L}_j$, we would conclude that ∇P_j was identically zero, leading to a contradiction. Therefore, $Z(P_j)$ contains a regular point.

But now the plane detection lemma says that $Z(P_j)$ is a plane for every j. The number of different factors P_j is $\leq \operatorname{Deg} P \lesssim L/A$. So we conclude that $\mathfrak{L}$ lies in $\lesssim L/A$ planes.

This finishes the proof of the planar clustering theorem.

11.8. Exercises

Exercise 11.1. Suppose that l_i are lines in $\mathbb{F}_q^3$ and that $X_i \subset l_i$ are subsets with $|X_i| \geq q/2$. Using the methods from the degree reduction argument, prove that

$$|\cup_i l_i| \lesssim (\log q)|\cup_i X_i|.$$

Hint: It is not hard to reduce to the special case that for every i, at least $q/4$ points of X_i also lie in $X_{i'}$ for some $i \neq i'$. In this special case, use the contagious vanishing argument to estimate the degree of $\cup_i X_i$. Prove that

$$\operatorname{Deg}(\cup_i X_i) \lesssim (\log q) q^{-2} |\cup_i X_i|.$$

But for any set $X \subset \mathbb{F}_q^3$, it is not hard to check that $|X| \lesssim q^2 \operatorname{Deg} X$. In fact, the Schwarz-Zippel lemma (Exercise 2.3) gives the estimate $|X| \leq q^2 \operatorname{Deg} X$.

If you work harder, it may be possible to remove the factor $\log q$. It seems to be quite hard to generalize this method to higher dimensions. See [**NW**] for a very different approach that works in all dimensions and avoids any factor of $\log q$.

Exercise 11.2. Suppose that $\mathbb{F}$ is an infinite field and that $N \geq \operatorname{Dim} \operatorname{Poly}_D(\mathbb{F}^n)$. Prove that there exists a set of N points in $\mathbb{F}^n$ with degree greater than D.

Exercise 11.3. Using Theorem 11.1, prove the following conjecture, made by Bourgain in [**CrLe**]. Suppose that $\mathfrak{L}$ is a set of N^2 lines in $\mathbb{R}^3$, with at most N lines in any plane. Suppose that $X \subset \mathbb{R}^3$ is a finite set. Suppose that each line $l \in \mathfrak{L}$ contains at least N points of X. Prove that $|X| \gtrsim N^3$.

EXERCISE 11.4. (*) Generalize all the arguments in the chapter to complex lines in $\mathbb{C}^3$. The only part of the argument that needs some modification is the discussion of flat points in Section 11.6. Conclude that Theorem 11.1 holds for lines in $\mathbb{C}^3$. As a corollary, show that the result from Exercise 11.3 also holds in $\mathbb{C}^3$: If $\mathfrak{L}$ is a set of N^2 complex lines in $\mathbb{C}^3$ with at most N lines in any complex 2-plane in $\mathbb{C}^3$, and if X is a finite set with at least N points of X on every line of $\mathfrak{L}$, then $|X| \gtrsim N^3$.

This result contrasts with an example using thin tubular neighborhoods of complex lines in Section 15.9.

EXERCISE 11.5. (*) Explore degree reduction for a set of lines in $\mathbb{F}^4$.

Suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{F}^4$. Prove that $\mathrm{Deg}(\mathfrak{L}) \lesssim L^{1/3}$.

Now suppose in addition that each line of $\mathfrak{L}$ contains at least A points of $P_2(\mathfrak{L})$. Prove that $\mathrm{Deg}(\mathfrak{L}) \lesssim L^{1/2}A^{-1/2}$. If A is much larger than $L^{1/3}$, then this bound improves on the bound from the last paragraph, which holds for an arbitrary set of lines in $\mathbb{F}^3$.

Here is a new twist in $\mathbb{F}^4$ as opposed to $\mathbb{F}^3$. If A is close to L, then we have shown that the lines of $\mathfrak{L}$ lie in a low degree 3-dimensional variety. But if A is close to L, we could hope to prove something stronger: the lines of $\mathfrak{L}$ may lie in a low degree 2-dimensional variety. The following paragraph gives an estimate in this spirit.

Suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{F}^4$, each line of $\mathfrak{L}$ contains at least A points of $P_2(\mathfrak{L})$, and all the lines of $\mathfrak{L}$ lie in $Z(P)$, the zero set of an irreducible polynomial P of degree $D \lesssim L^{1/2}A^{-1/2}$. Prove that there is a polynomial Q, which is not a multiple of P, so that the lines of $\mathfrak{L}$ lie in the zero set of Q as well, and with the degree bound $(\mathrm{Deg}\, P)(\mathrm{Deg}\, Q) \lesssim LA^{-1}$.

(**) Try to set up a theory of degree reduction for lines in $\mathbb{F}^n$ for all n. This is currently an open problem for $n \geq 5$.

https://doi.org/10.1090//ulect/064/12

CHAPTER 12

An incidence bound for lines in three dimensions

In this chapter, we combine the polynomial partitioning technique from Chapter 10 and the technique of flat points from Chapter 11. Using these methods together, we prove Theorem 8.4. We restate the theorem here:

THEOREM. If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with at most $L^{1/2}$ lines in any plane, and if $3 \leq r \leq 2L^{1/2}$, then $|P_r(\mathfrak{L})| \lesssim L^{3/2} r^{-2}$.

This theorem follows from an incidence bound for points and lines in $\mathbb{R}^3$, which one can think of as a 3-dimensional version of the Szemerédi-Trotter theorem.

THEOREM 12.1. Let $\mathcal{S}$ be a set of S points and $\mathfrak{L}$ a set of L lines in $\mathbb{R}^3$. Suppose that there are at most B lines of $\mathfrak{L}$ in any plane, and that $B \geq L^{1/2}$. Then the number of incidences is bounded as follows:

$$I(\mathcal{S}, \mathfrak{L}) \lesssim S^{1/2} L^{3/4} + B^{1/3} L^{1/3} S^{2/3} + L + S.$$

This bound on incidences gives a bound on r-rich points for all sufficiently large r.

COROLLARY 12.2. There is some constant $r_0 > 0$ so that the following holds. If $r \geq r_0$, and if $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with $\leq B$ in any plane, and if $B \geq L^{1/2}$, then

$$|P_r(\mathfrak{L})| \lesssim L^{3/2} r^{-2} + LBr^{-3} + Lr^{-1}.$$

PROOF. We apply the incidence estimate from Theorem 12.1. We let $\mathcal{S} = P_r(\mathfrak{L})$. We have

$$rS \leq I(\mathcal{S}, \mathfrak{L}) \lesssim S^{1/2} L^{3/4} + B^{1/3} L^{1/3} S^{2/3} + L + S.$$

Therefore, one of the following holds:

(1) $rS \leq C S^{1/2} L^{3/4}$.
(2) $rS \leq C B^{1/3} L^{1/3} S^{2/3}$.
(3) $rS \leq CL$.
(4) $rS \leq CS$.

Option (1) gives $S \lesssim L^{3/2} r^{-2}$. Option (2) gives $S \lesssim BLr^{-3}$. Option (3) gives $S \lesssim Lr^{-1}$. Option (4) gives $r \leq C$. If we choose $r_0 > C$, then one of Options (1) - (3) must occur, and we get the desired bound. □

In Chapter 11, we proved good estimates for 3-rich points. In particular, Corollary 11.8 gives the following bound:

COROLLARY. Suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ that contains at most B lines in any plane. If $B \geq L^{1/2}$, then

$$|P_3(\mathfrak{L})| \lesssim BL.$$

Plugging in $B = L^{1/2}$ in Corollary 12.2 and Corollary 11.8 gives Theorem 8.4.

The incidence bound in Theorem 12.1 is sharp up to a constant factor (for any S and L and any B in the range $L^{1/2} \leq B \leq L$). There are three types of examples to show that the bound is sharp. The upper bound on the number of incidences is $S^{1/2}L^{3/4} + B^{1/3}L^{1/3}S^{2/3} + L + S$. To get S incidences, choose all the points on a single line. To get L incidences, choose all the lines through a single point. To get $B^{1/3}L^{1/3}S^{2/3}$ incidences, choose LB^{-1} planes Π_i. For each plane Π_i, let $\mathfrak{L}_i$ be a set of B lines in Π_i and let $\mathcal{S}_i$ be a set of SBL^{-1} points in Π_i. Choose $\mathfrak{L}_i$ and $\mathcal{S}_i$ in a grid example, so that they are optimal for the Szemerédi-Trotter theorem. Then the number of incidences between $\mathfrak{L}_i$ and $\mathcal{S}_i$ is $\gtrsim B^{2/3}(SBL^{-1})^{2/3}$. Therefore the total number of incidences is $\gtrsim LB^{-1} \cdot B^{2/3}(SBL^{-1})^{2/3} = B^{1/3}L^{1/3}S^{2/3}$. To get $S^{1/2}L^{3/4}$ incidences, use the example in Exercise 8.1. This example is a set of L lines in $\mathbb{R}^3$ with at most $L^{1/2}$ lines of $\mathfrak{L}$ in any plane, and it has $\sim L^{3/2}r^2$ r-rich points for any $2 \leq r \leq L^{1/2}/400$. Setting $\mathcal{S} = P_r(\mathfrak{L})$, we see that the number of incidences is

$$|I(\mathfrak{L}, \mathcal{S})| \geq r|P_r(\mathfrak{L})| \sim L^{3/2}r^{-1} = (L^{3/2}r^{-2})^{1/2}L^{3/4} = S^{1/2}L^{3/4}.$$

In the rest of the chapter we prove Theorem 12.1. The proof involves the polynomial partitioning method from Chapter 10 and the technique of flat points from Chapter 11. There aren't really new ideas in the proof, but the organization is a little bit complex, with many different terms. Therefore, as a warmup, we give a slightly different proof of the Szemerédi-Trotter theorem following a similar outline.

12.1. Warmup: The Szemerédi-Trotter theorem revisited

We begin by reproving the Szemerédi-Trotter theorem using polynomial partitioning, organizing the proof in a way that will help get ready for the 3-dimensional version in Theorem 12.1.

THEOREM 12.3. (Szemerédi-Trotter) If $\mathcal{S}$ is a set of S points and $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^2$, then the number of incidences obeys the following bound:

$$I(\mathcal{S}, \mathfrak{L}) \lesssim S^{2/3}L^{2/3} + S + L.$$

We will prove the result by using a polynomial cell decomposition together with elementary counting bounds in each cell. We first recall the counting bounds.

LEMMA 12.4. If $\mathcal{S}$ and $\mathfrak{L}$ are as above, then

- $I(\mathcal{S}, \mathfrak{L}) \leq L + S^2$.
- $I(\mathcal{S}, \mathfrak{L}) \leq L^2 + S$.

PROOF. Fix $x \in \mathcal{S}$. Let L_x be the number of lines of $\mathfrak{L}$ that contain x and no other point of $\mathcal{S}$. For each other point $y \in S$, there is at most one line of $\mathfrak{L}$ containing x and y. Therefore, $I(x, \mathfrak{L}) \leq S + L_x$. So $I(\mathcal{S}, \mathfrak{L}) \leq S^2 + \sum_{x \in S} L_x \leq S^2 + L$.

The proof of the other inequality is similar. □

Now we turn to the proof of the theorem.

PROOF. If $L > S^2/10$ or $S > L^2/10$, then the conclusion follows from Lemma 12.4. Therefore, we can now restrict to the case that

$$10^{1/2}S^{1/2} \leq L \leq S^2/10. \tag{1}$$

We will also use induction on L, and so we can assume the theorem holds for smaller sets of lines. More precisely, we let C_0 be a large constant to choose later. We want to prove that

$$I(\mathcal{S}, \mathfrak{L}) \leq C_0 \left(S^{2/3} L^{2/3} + S + L \right).$$

We can assume that this inequality holds for any set of at most $L/2$ lines, and we have to prove that it also holds for a set of L lines. (For the base of the induction, we take $L = 1$, and then the number of incidences is clearly at most S.)

Now we come to the heart of the proof. We use the polynomial cell decomposition to cut $\mathbb{R}^2$ into cells, and then we use the counting lemma in each cell.

Let D be a degree to choose later. By the polynomial partitioning theorem, Theorem 10.3, we can find a non-zero polynomial P of degree $\leq D$ so that each component of the complement of $Z(P)$ contains $\lesssim SD^{-2}$ points of $\mathcal{S}$. Let O_i be the components, S_i the number of points of $\mathcal{S}$ in O_i, and L_i the number of lines of $\mathfrak{L}$ that intersect O_i. Since each line intersects $\leq D + 1$ cells, we know that $\sum L_i \leq L(D + 1)$.

Applying the counting lemma in each cell, we get

$$I(\mathcal{S}_i, \mathfrak{L}_i) \leq L_i + S_i^2.$$

We let $\mathcal{S}_{cell}$ be the union of $\mathcal{S}_i$ - all the points of $\mathcal{S}$ that lie in the interiors of the cells.

$$I(\mathcal{S}_{cell}, \mathfrak{L}) = \sum_i I(\mathcal{S}_i, \mathfrak{L}_i) \leq \sum_i L_i + \sum_i S_i^2 \lesssim LD + SD^{-2} \sum_i S_i \leq LD + S^2 D^{-2}.$$

We let $\mathcal{S} = \mathcal{S}_{cell} \cup \mathcal{S}_{alg}$, where $\mathcal{S}_{alg}$ is the set of points in $Z(P)$. It remains to bound $I(\mathcal{S}_{alg}, \mathfrak{L})$. We divide $\mathfrak{L}$ as $\mathfrak{L}_{cell} \cup \mathfrak{L}_{alg}$, where $\mathfrak{L}_{cell}$ are the lines that intersect some open cells, and $\mathfrak{L}_{alg}$ are the lines contained in $Z(P)$.

Each line of $\mathfrak{L}_{cell}$ has $\leq D$ intersection points with $Z(P)$, hence $\leq D$ incidences with $\mathcal{S}_{alg}$. Hence $I(\mathcal{S}_{alg}, \mathfrak{L}_{cell}) \leq LD$. Summarizing everything so far, we have the following:

$$I(\mathcal{S}, \mathfrak{L}) \leq C(LD + S^2 D^{-2}) + I(\mathcal{S}_{alg}, \mathfrak{L}_{alg}).$$

We will deal with the last term by induction. We will choose $D \leq L/2$. So $\mathfrak{L}_{alg}$ contains $\leq L/2$ lines. By induction,

$$I(\mathcal{S}_{alg}, \mathfrak{L}_{alg}) \leq C_0 [S^{2/3} (L/2)^{2/3} + S + L/2].$$

Now we are ready to optimize over D. We need to choose D to be an integer between 1 and $L/2$. We choose $dD \sim S^{2/3} L^{-1/3}$. Because of the bounds in equation (1), we can find D this size in the range $1 \leq D \leq L/2$. Plugging in, we get

$$I(\mathcal{S}, \mathfrak{L}) \leq C L^{2/3} S^{2/3} + C_0 [S^{2/3} (L/2)^{2/3} + S + L/2].$$

Finally, we choose C_0 large enough compared to C, and the whole right hand side is bounded by $C_0 [S^{2/3} L^{2/3} + S + L]$. □

12.2. Three-dimensional incidence estimates

In this section, we prove Theorem 12.1. We restate it here for convenience:

THEOREM. There is a large constant C_0 so that the following holds. Let $\mathcal{S}$ be a set of S points and $\mathfrak{L}$ a set of L lines in $\mathbb{R}^3$. Suppose that there are at most B lines of $\mathfrak{L}$ in any plane, and that $B \geq L^{1/2}$. Then the number of incidences is bounded as follows:

$$I(\mathcal{S}, \mathfrak{L}) \leq C_0 \left[S^{1/2} L^{3/4} + B^{1/3} L^{1/3} S^{2/3} + L + S \right]. \qquad (*)$$

The proof follows a similar outline to the proof of Szemerédi-Trotter in the last section. Unfortunately, there are many different terms, making the argument complicated. I apologize to the reader for this complexity. We do a polynomial cell decomposition with a polynomial $Z(P)$. There are three main contributions. We use the polynomial partitioning theorem to control the incidences in the cells (outside of $Z(P)$). We divide the surface $Z(P)$ into planar parts and non-planar parts. The contribution from the planar parts is controlled using the fact that there are at most B lines in any plane. The contribution from the non-planar parts of $Z(P)$ is controlled using the theory of flat points and lines. When we carry out this argument, there end up being a lot of terms. Most of the terms fall into the three main contributions descibed above. There are also some contributions that involve small numbers of lines, which we handle by induction on L, as we saw in the last section.

PROOF. The proof is by induction on L. We assume that $(*)$ holds for sets of $< L$ lines, and we want to prove that it also holds for $\mathfrak{L}$.

We have some previous estimates for $I(\mathcal{S}, \mathfrak{L})$. The counting argument in Lemma 12.4 gives

$$I(\mathcal{S}, \mathfrak{L}) \leq L + S^2; I(\mathcal{S}, \mathfrak{L}) \leq L + S^2.$$

Also the Szemerédi-Trotter theorem applies in any dimension by a random projection argument (cf. Proposition 8.1), giving the estimate

$$I(\mathcal{S}, \mathfrak{L}) \lesssim \left[S^{2/3} L^{2/3} + L + S \right].$$

Because of the counting bounds, we can assume that $10L^{1/2} \leq S \leq (1/10)L^2$.

Let $D \geq 1$ be an integer that we will choose later. We do a polynomial cell decomposition of degree D for the point set $\mathcal{S}$. By the polynomial partitioning theorem, Theorem 10.3, we can choose a non-zero polynomial P of degree $\leq D$ so that each component of $\mathbb{R}^3 \setminus Z(P)$ contains $\lesssim SD^{-3}$ points of $\mathcal{S}$.

First we estimate the incidences coming from points outside of $Z(P)$. We make some vocabulary to describe which objects are in $Z(P)$:

- $\mathcal{S}_{alg}$ is the set of points of $\mathcal{S}$ that lie in $Z(P)$.
- $\mathcal{S}_{cell} = \mathcal{S} \setminus \mathcal{S}_{alg}$.
- $\mathfrak{L}_{alg}$ is the set of lines of $\mathfrak{L}$ that lie in $Z(P)$.
- $\mathfrak{L}_{cell} = \mathfrak{L} \setminus \mathfrak{L}_{alg}$.

LEMMA 12.5. (Cellular estimate) For some constant C,

$$I(\mathcal{S}, \mathfrak{L}) \leq C \left[D^{-1/3} L^{2/3} S^{2/3} + DL + S_{cell} \right] + I(\mathcal{S}_{alg}, \mathfrak{L}_{alg}).$$

PROOF. Let O_i be the components of $\mathbb{R}^3 \setminus Z(P)$. We define

- $\mathcal{S}_i = \mathcal{S} \cap O_i$.
- $\mathfrak{L}_i$ is the set of lines of $\mathfrak{L}$ that intersect O_i.

We know that $\sum S_i = S_{cell}$, that $S_i \lesssim SD^{-3}$, and that $\sum_i L_i \lesssim DL$.

Now $I(\mathcal{S}_{cell}, \mathfrak{L}) = \sum_i I(\mathcal{S}_i, \mathfrak{L}_i)$. We bound each term of the sum using Szemerédi-Trotter.

$$\sum_i I(\mathcal{S}_i, \mathfrak{L}_i) \lesssim \sum_i L_i^{2/3} S_i^{2/3} + L_i + S_i \lesssim DL + S_{cell} + \sum_i L_i^{2/3} S_i^{2/3}.$$

To bound the last term, we recall that $S_i \lesssim SD^{-3}$ and then apply Holder:

$$\sum_i L_i^{2/3} S_i^{2/3} \lesssim S^{1/3} D^{-1} \sum_i L_i^{2/3} S_i^{1/3} \leq S^{1/3} D^{-1/3} (\sum_i L_i)^{2/3} (\sum_i S_i)^{1/3} \lesssim$$

$$\lesssim S^{1/3} D^{-1} (DL)^{2/3} S^{1/3} = D^{-1/3} S^{2/3} L^{2/3}.$$

Altogether, we have

$$I(\mathcal{S}_{cell}, \mathfrak{L}) \lesssim D^{-1/3} S^{2/3} L^{2/3} + L + S_{cell}. \tag{12.1}$$

On the other hand, each line of $\mathfrak{L}_{cell}$ intersects at most D points of $\mathcal{S}_{alg}$, and so

$$I(\mathcal{S}_{alg}, \mathfrak{L}_{cell}) \leq DL. \tag{12.2}$$

We know $I(\mathcal{S}, \mathfrak{L}) \leq I(\mathcal{S}_{cell}, \mathfrak{L}) + I(\mathcal{S}_{alg}, \mathfrak{L}_{cell}) + I(\mathcal{S}_{alg}, \mathfrak{L}_{alg})$. Combining inequalities 12.1 and 12.2, we get the conclusion of this lemma. □

Our next goal is to control the algebraic incidences $I(\mathcal{S}_{alg}, \mathfrak{L}_{alg})$. For the argument in the 2-dimensional case, in Section 12.1, we knew that $|\mathfrak{L}_{alg}| \leq D$, and we chose D to arrange $|\mathfrak{L}_{alg}| \leq L/2$, allowing us to handle this term by induction. In the 3-dimensional case, there is no bound on the size of $\mathfrak{L}_{alg}$, and we have to pay more attention to this term.

Some of the incidences in $I(\mathcal{S}_{alg}, \mathfrak{L}_{alg})$ may come from points and lines that lie in planes of $Z(P)$. To control these incidences, we need to use the hypothesis that $\leq B$ lines of $\mathfrak{L}$ lie in any plane. We make the following vocabulary to describe the contribution of the planes in $Z(P)$.

- $\mathfrak{L}_{plan}$ is the set of lines of $\mathfrak{L}$ contained in at least one plane in $Z(P)$.
- $\mathfrak{L}_{uniplan}$ is the set of lines of $\mathfrak{L}$ contained in exactly one plane in $Z(P)$.
- $\mathfrak{L}_{multiplan}$ is the set of lines of $\mathfrak{L}$ contained in at least two planes in $Z(P)$.

Similarly, we can define subsets of $\mathcal{S}$:

- $\mathcal{S}_{plan}$ is the set of points of $\mathcal{S}$ contained in at least one plane in $Z(P)$.
- $\mathcal{S}_{uniplan}$ is the set of points of $\mathcal{S}$ contained in exactly one plane in $Z(P)$.
- $\mathcal{S}_{multiplan}$ is the set of points of $\mathcal{S}$ contained in at least two planes in $Z(P)$.

We prove the following estimate for incidences involving planes.

LEMMA 12.6. (Planar estimate)

$$I(\mathcal{S}_{alg}, \mathfrak{L}_{plan}) \leq C \left(B^{1/3} L^{1/3} S^{2/3} + DL + S_{uniplan} \right) + I(\mathcal{S}_{multplan}, \mathfrak{L}_{multplan}).$$

The statement is a little complicated and so we discuss it. When we eventually choose D, we will choose it so that the first term on the right hand side is acceptable for $(*)$. There cannot be too many lines in $\mathfrak{L}_{multplan}$, and we will handle the contribution of $\mathfrak{L}_{multplan}$ by induction. Since $Z(P)$ contains at most D planes, $|\mathfrak{L}_{multplan}| \leq \binom{D}{2} \leq D^2$. We note this inequality for later:

$$|\mathfrak{L}_{multplan}| \leq D^2. \tag{12.3}$$

We will choose D so that $D^2 \leq L/10$, and then we will handle the contribution of $\mathfrak{L}_{multplan}$ by induction on L. Now we turn to the proof of Lemma 12.6

Proof. We first observe that

$$I(\mathcal{S}_{alg}, \mathfrak{L}_{plan}) = I(\mathcal{S}_{alg}, \mathfrak{L}_{uniplan}) + I(\mathcal{S}_{multplan}, \mathfrak{L}_{multplan}).$$

Therefore, it suffices to prove that

$$I(\mathcal{S}_{alg}, \mathfrak{L}_{uniplan}) \lesssim B^{1/3} L^{1/3} S^{2/3} + DL + S_{uniplan}.$$

Morally, we just use Szemerédi-Trotter in each plane of $Z(P)$ and add up the results. If π is a plane in $Z(P)$, we let $\mathfrak{L}_\pi \subset \mathfrak{L}_{plan}$ be the set of lines of $\mathfrak{L}$ that lie in π, and we let $\mathfrak{L}_{uni\pi} \subset \mathfrak{L}_{uniplan}$ be the set of lines of $\mathfrak{L}$ that lie in π and in no other plane of $Z(P)$. We define $\mathcal{S}_\pi$ and $\mathcal{S}_{uni\pi}$ in the same way. Next we observe that

$$I(\mathcal{S}_{alg}, \mathfrak{L}_{uniplan}) = \sum_\pi I(\mathcal{S}_\pi, \mathfrak{L}_{uni\pi}).$$

If l is a line of $\mathfrak{L}_{uni\pi}$, the l can contain at most D points of $\mathcal{S}_{multplan}$. Therefore,

$$\sum_\pi I(\mathcal{S}_\pi, \mathfrak{L}_{uni\pi}) \leq DL + \sum_\pi I(\mathcal{S}_{uni\pi}, \mathfrak{L}_{uni\pi}).$$

The sets $S_{uni\pi}$ are disjoint (as π varies), and so $\sum_\pi S_{uni\pi} \leq S_{uniplan}$. Similarly, $\sum_\pi L_{uni\pi} \leq L$. Now we bound the last sum by applying Szemerédi-Trotter to each term:

$$\begin{aligned}\sum_\pi I(\mathcal{S}_{uni\pi}, \mathfrak{L}_{uni\pi}) &\lesssim \sum_\pi (L_{uni\pi}^{2/3} S_{uni\pi}^{2/3} + L_{uni\pi} + S_{uni\pi}) \\ &\leq L + S_{uniplan} + \sum_\pi L_{uni\pi}^{2/3} S_{uni\pi}^{2/3}.\end{aligned}$$

To bound the last sum, we use that $L_{uni\pi} \leq B$ and apply Holder:

$$\begin{aligned}\sum_\pi L_{uni\pi}^{2/3} S_{uni\pi}^{2/3} &\leq B^{1/3} \sum_\pi L_{uni\pi}^{1/3} S_{uni\pi}^{2/3} \\ &\leq B^{1/3} (\sum_\pi L_{uni\pi})^{1/3} (\sum_\pi S_{uni\pi})^{2/3} \leq B^{1/3} L^{1/3} S^{2/3}.\end{aligned}$$

Assembling our estimates we get the desired conclusion:

$$I(\mathcal{S}_{alg}, \mathfrak{L}_{uniplan}) \lesssim B^{1/3} L^{1/3} S^{2/3} + DL + S_{uniplan}.$$

□

We have now bounded the incidences coming from points and lines in the planar part of $Z(P)$. Next we turn to the points and lines in the rest of $Z(P)$. These bounds involve the theory of critical and flat points in $Z(P)$ which we studied in Section 11.6.

Recall that we say a point $x \in Z(P)$ is special if x is critical or flat. We say that a line $l \subset Z(P)$ is special if each point of the line is special. Now we define some subsets of $\mathcal{S}$ and $\mathfrak{L}$ that have to do with special points and lines.

- $\mathcal{S}_{spec}$ is the subset of $\mathcal{S}$ consisting of special points of $Z(P)$.
- $\mathcal{S}_{nonspec} = \mathcal{S}_{alg} \setminus \mathcal{S}_{spec}$
- $\mathfrak{L}_{spec}$ is the subset of $\mathfrak{L}$ consisting of special lines of $Z(P)$.
- $\mathfrak{L}_{nonspec} = \mathfrak{L}_{alg} \setminus \mathfrak{L}_{spec}$.

In Section 10.5 we proved the following result about special points and lines:

Plane detection lemma. For any polynomial P in $\mathbb{R}[x_1, x_2, x_3]$, we can associate a list of polynomials SP with the following properties.

(1) If $x \in Z(P)$ then $SP(x) = 0$ iff x is critical or flat.
(2) If x is contained in three lines in $Z(P)$, then $SP(x) = 0$.
(3) $\operatorname{Deg} SP \leq 3 \operatorname{Deg} P$.
(4) If P is irreducible and SP vanishes on $Z(P)$ and $Z(P)$ contains a regular point, then $Z(P)$ is a plane.

This result allows us to bound the contribution from non-special lines in $Z(P)$. More precisely:

Lemma 12.7. (Algebraic estimate)

$$I(\mathcal{S}_{alg}, \mathfrak{L}_{alg} \setminus \mathfrak{L}_{plan}) \leq C(DL + S_{nonspec}) + I(\mathcal{S}_{spec}, \mathfrak{L}_{spec} \setminus \mathfrak{L}_{plan}).$$

Proof. We note that

$$I(\mathcal{S}_{alg}, \mathfrak{L}_{alg} \setminus \mathfrak{L}_{plan}) \leq I(\mathcal{S}_{nonspec}, \mathfrak{L}_{alg}) + I(\mathcal{S}_{spec}, \mathfrak{L}_{nonspec}) + I(\mathcal{S}_{spec}, \mathfrak{L}_{spec} \setminus \mathfrak{L}_{plan}).$$

By item (2) of the plane detection lemma, $I(\mathcal{S}_{nonspec}, \mathfrak{L}_{alg}) \leq 2S_{nonspec}$.
By item (3) of the plane detection lemma, $I(\mathcal{S}_{spec}, \mathfrak{L}_{nonspec}) \leq 3DL$.

□

We still have to control the contribution of the special lines – more precisely we have to bound $I(\mathcal{S}_{spec}, \mathfrak{L}_{spec} \setminus \mathfrak{L}_{plan})$. The last key point is that there are few lines in $\mathfrak{L}_{spec} \setminus \mathfrak{L}_{plan}$, allowing us to control this term by induction. In particular, we will prove the following bound on $|\mathfrak{L}_{spec} \setminus \mathfrak{L}_{plan}|$:

$$|\mathfrak{L}_{spec} \setminus \mathfrak{L}_{plan}| \leq 4D^2. \tag{12.4}$$

This estimate follows from a bound on the number of special lines in an algebraic surface. We begin with the irreducible case.

Proposition 12.8. If P is irreducible and $Z(P)$ is not a plane, then $Z(P)$ contains $\leq 3(\operatorname{Deg} P)^2$ special lines.

Proof. Suppose that $Z(P)$ has a regular point. If SP vanished on $Z(P)$, then the plane detection lemma would imply that $Z(P)$ was a plane. Therefore, SP does not vanish on $Z(P)$. Let Q be one of the polynomials in the list SP that does not vanish on $Z(P)$. Since P is irreducible, P and Q have no common factor. Note that SP vanishes on each special line, so the special lines lie in $Z(P) \cap Z(Q)$. But by the Bezout theorem for lines, Theorem 6.7, $Z(P) \cap Z(Q)$ contains at most $3(\operatorname{Deg} P)^2$ lines.

Now suppose that $Z(P)$ has no regular point. Then $\partial_i P$ vanishes on $Z(P)$ for each i. Since P is not constant, we can choose i so that $\partial_i P$ is not the zero polynomial. Since P is irreducible and $\operatorname{Deg} \partial_i P < \operatorname{Deg} P$, we see that P and $\partial_i P$

have no common factor. Using the Bezout theorem for lines, Theorem 6.7, we see that $Z(P) \subset Z(P) \cap Z(\partial_i P)$ contains at most D^2 lines. □

Now we turn to the general case.

PROPOSITION 12.9. If P is any (square-free) non-zero polynomial, then there are at most $4(\mathrm{Deg}\, P)^2$ special lines of $Z(P)$ that are not contained in any plane in $Z(P)$.

This Proposition implies inequality 12.4.

PROOF. Suppose that $P = \prod P_j$, where P_j are irreducible and distinct. We claim that a line $l \subset Z(P)$ is a special line for P if and only if either l is a special line of P_j for some j or l lies in $Z(P_j)$ for more than one j. Assuming the claim for a moment, let us count special lines of $Z(P_j)$. By Bezout, Theorem 6.7, the number of lines lying in $Z(P_j)$ for more than one j is $\leq (\mathrm{Deg}\, P)^2$. The number of special lines in $Z(P_j)$ is $\leq 3(\mathrm{Deg}\, P_j)^2$. The total number of special lines in all $Z(P_j)$ is $\leq \sum_j 3(\mathrm{Deg}\, P_j)^2 \leq 3(\mathrm{Deg}\, P)^2$. So the total number of special lines in $Z(P)$ is $\leq 4(\mathrm{Deg}\, P)^2$.

Now we prove the claim. First suppose that a line l is contained in $Z(P_i)$ and $Z(P_j)$ for $i \neq j$. At any point $x \in l$, $\nabla P(x) = 0$. To see this, expand $\nabla P = \sum_k (\nabla P_k) P_1 ... P_{k-1} P_{k+1} ...$ and note that each term of the sum vanishes at x. So every point of l is critical and l is a special line. Now suppose that $l \subset Z(P_j)$ for a unique j. Along l, we have $\nabla P = (\nabla P_j) P_1 ... P_{j-1} P_{j+1} ....$ Since P_i $i \neq j$ vanishes at only finitely many points of l, we see that ∇P vanishes on l if and only if ∇P_j vanishes on l. So l is a critical line of $Z(P)$ if and only if l is a critical line of $Z(P_j)$. Finally suppose that l is not a critical line of $Z(P)$ or $Z(P_j)$. Then l is special if and only if every regular point of l is flat. But we can check flatness of $Z(P)$ near a regular point x by examining $Z(P)$ in a small neighborhood of x. For a regular point x, there is a small neighborhood $x \in U$ where $Z(P) \cap U = Z(P_j) \cap U$. So a regular point x is flat for $Z(P)$ if and only if x is flat for $Z(P_j)$. This proves the claim.

□

When we put together the three lemmas, we get an estimate for lots of the incidences in $I(\mathcal{S}, \mathfrak{L})$ plus a leftover term involving special and multiplanar lines. We will bound the leftover term by induction. We let $\mathfrak{L}_{leftover} = \mathfrak{L}_{multplan} \cup (\mathfrak{L}_{spec} \setminus \mathfrak{L}_{plan})$, and we define $\mathcal{S}_{leftover} = \mathcal{S}_{multplan} \cup \mathcal{S}_{spec}$. We define $\mathcal{S}_{main}$ to be $\mathcal{S} \setminus \mathcal{S}_{leftover}$. Assembling our estimates, we get the following:

$$I(\mathcal{S}, \mathfrak{L}) \leq C \left[D^{-1/3} S^{2/3} L^{2/3} + DL + B^{1/3} L^{1/3} S^{2/3} + L + S_{main} \right] +$$

$$+ I(\mathcal{S}_{leftover}, \mathfrak{L}_{leftover}).$$

Moreover, we have proven that $|\mathfrak{L}_{multplan}| \leq D^2$ (inequality 12.3) and $|\mathfrak{L}_{spec} \setminus \mathfrak{L}_{plan}| \leq 4D^2$ (inequality 12.4), and so

$$|\mathfrak{L}_{leftover}| \leq 10 D^2.$$

Now we choose D in the range $1 \leq D \leq (1/10) L^{1/2}$ in order to minimize the term in brackets above. Since $D \leq (1/10) L^{1/2}$, we have $|\mathfrak{L}_{leftover}| \leq L/2$, allowing us to apply induction to the leftover term. The rest of the proof is just a calculation.

When we choose the optimal value of D in the range $1 \leq D \leq (1/10)L^{1/2}$, we claim that we will get:

$$D^{-1/3}S^{2/3}L^{2/3} + DL \lesssim S^{1/2}L^{3/4} + B^{1/3}L^{1/3}S^{2/3}. \tag{12.5}$$

We will check this claim below by computation. Given the claim, and given that $|\mathfrak{L}_{leftover}| \leq 10D^2 \leq L/2$, we see by induction that

$$I(\mathcal{S}, \mathfrak{L}) \leq C\left[S^{1/2}L^{3/4} + B^{1/3}L^{1/3}S^{2/3} + S_{main} + L\right] +$$

$$+C_0\left[S^{1/2}(L/2)^{3/4} + B^{1/3}(L/2)^{1/3}S^{2/3} + (L/2) + S_{leftover}\right].$$

At this point, we choose C_0 sufficiently large compared to C, and we get the desired inequality $(*)$.

Finally we check the claim that we can choose D in the range $1 \leq D \leq (1/10)L^{1/2}$, so that inequality 12.5 holds.

To minimize $D^{-1/3}S^{2/3}L^{2/3} + DL$, we want to choose D to balance the two terms: $D^{-1/3}S^{2/3}L^{2/3}$ with DL. The balancing is achieved by setting $D \sim S^{1/2}L^{-1/4}$. Because of the counting estimates, we have been able to assume from the beginning that $10L^{1/2} \leq S \leq (1/10)L^2$. This implies that $1 \leq S^{1/2}L^{-1/4} \leq L^{3/4}$. However, in order to apply the induction, we need to choose D in the range $1 \leq D \leq (1/10)L^{1/2}$. There are now two cases depending on whether $S^{1/2}L^{-1/4}$ is larger than $(1/10)L^{1/2}$.

If $S^{1/2}L^{-1/4} \leq (1/10)L^{1/2}$, then we set D to be (the greatest integer at most) $S^{1/2}L^{-1/4}$, and $D^{-1/3}S^{2/3}L^{2/3} + DL \sim S^{1/2}L^{3/4}$.

If $S^{1/2}L^{-1/4} > (1/10)L^{1/2}$, the we set D to be (the greatest integer at most) $(1/10)L^{1/2}$. In this case, $D^{-1/3}S^{2/3}L^{2/3} + DL$ is dominated by $D^{-1/3}S^{2/3}L^{2/3} \sim S^{2/3}L^{1/2}$. But since $B \geq L^{1/2}$, $S^{2/3}L^{1/2} \leq B^{1/3}L^{1/3}S^{2/3}$. This checks inequality 12.5. □

https://doi.org/10.1090//ulect/064/13

CHAPTER 13

Ruled surfaces and projection theory

In this chapter, we continue the theme of connecting combinatorial structure and algebraic structure. In terms of combinatorics, we will study 2-rich points of a collection of lines. The main result of this Chapter is Theorem 8.3, which we restate here for convenience.

THEOREM. If $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with at most B lines in any plane or degree 2 surface, then $|P_2(\mathfrak{L})| \lesssim LB + L^{3/2}$.

More generally, we will prove that this inequality holds for lines in $\mathbb{C}^3$:

THEOREM 13.1. If $\mathfrak{L}$ is a set of L lines in $\mathbb{C}^3$ with at most B lines in any plane or degree 2 surface, then $|P_2(\mathfrak{L})| \lesssim LB + L^{3/2}$.

Recall that in Section 3.5, we saw an example where all the lines of $\mathfrak{L}$ lie in the degree 2 surface $z = xy$, and where $\mathfrak{L}$ has $\sim L^2$ 2-rich points. (This degree 2 surface was an example of a regulus, which we discussed more in Section 8.4.) So we really do need to mention degree 2 surfaces in the statement of the theorem. The presence of degree 2 surfaces and not just planes makes this theorem a lot more complex and it brings in some interesting ideas from algebraic geometry.

The arguments in this chapter involve the theory of ruled surfaces from algebraic geometry. We begin this Chapter with a long introduction, introducing ruled surface theory and explaining how it becomes relevant to our combinatorial problem.

Recall that an algebraic surface $Z(P)$ is called ruled if every point of $Z(P)$ lies in a line of $Z(P)$. An algebraic surface is called doubly ruled if every point of $Z(P)$ lies in two distinct lines in $Z(P)$. The surface $z = xy$ is doubly ruled, as we saw in Section 3.5. Based on this example, we might try to build configurations of lines with many 2-rich points by taking lines inside of other doubly ruled surfaces. There is a classification of doubly ruled surfaces in $\mathbb{C}^3$, which I believe was proven in the 19th century.

THEOREM 13.2. Suppose that $P \in \mathrm{Poly}(\mathbb{C}^3)$ is an irreducible polynomial, and suppose that $Z(P)$ is doubly ruled. Then P has degree 1 or 2, and $Z(P)$ is a plane or regulus.

This theorem is mostly based on the theory of reguli from Section 8.4. We will prove it later in the chapter. We can think of Theorem 13.1 as a strong generalization of the classification of doubly ruled surfaces. Not only is it true that an irreducible algebraic surface with two lines through every point is a plane or a regulus, but we will see that a sufficiently big finite configuration of lines with many points lying in two lines must be modeled on a small number of planes or reguli.

The proof in this chapter uses some of the ideas from Chapter 11, but it also requires some significant new ideas related to the theory of ruled surfaces. In

Chapter 11, we proved that if $\mathfrak{L}$ has at most B lines in any plane, then $|P_3(\mathfrak{L})| \lesssim BL + L^{3/2}$. We give a high-level summary of the proof so that we can introduce the new issues and ideas.

The proof was by contradiction: we assumed that $|P_3(\mathfrak{L})|$ was too large and then we proved that many lines of $\mathfrak{L}$ lie in a plane. We considered a lowest degree polynomial P that vanishes on the lines of $\mathfrak{L}$. By the degree reduction argument, Proposition 11.5, we got a strong bound on $\operatorname{Deg} P$. Then we showed that each point of $P_3(\mathfrak{L})$ is a special point of $Z(P)$: either a critical point or a flat point. Being a critical or flat point has an algebraic interpretation: a point $x \in Z(P)$ is critical or flat if and only if some polynomials $SP(x)$ vanish, and $\operatorname{Deg} SP \lesssim \operatorname{Deg} P$. Since there are many 3-rich points, and since $\operatorname{Deg} P$ is small, the critical or flat points are contagious, and we were able to prove that every point of $Z(P)$ is critical or flat. But if every point of $Z(P)$ is critical or flat, then $Z(P)$ must be a union of planes, and the number of planes is at most $\operatorname{Deg} P$. Since all the lines of $\mathfrak{L}$ lie in $Z(P)$, there must be a plane containing many lines.

When we switch from 3-rich points to 2-rich points, we encounter a new difficulty. Is a point of $P_2(\mathfrak{L})$ a special point of $Z(P)$? It does not need to be either critical or flat. What other special feature might it have? And if we find a special feature of these points, how do we know whether the special feature is contagious?

Our approach to this question comes from the theory of ruled surfaces. A point $z \in Z(P)$ is called flecnodal if there is a line through z where P vanishes to third order. In talking about flecnodes, it is helpful to introduce the jets of a polynomial. The r-jet $J^r P(z) \in \operatorname{Poly}_r(\mathbb{C}^3)$ is the r^{th} order Taylor series of P at z. In other words, $J^r P(z)$ is the unique polynomial of degree at most r so that

$$P(z+h) = J^r P(z)(h) + O(|h|^{r+1}). \tag{13.1}$$

(In this equation, h is a variable in $\mathbb{C}^3$, and $J^r P(z)(h)$ is the evaluation of the polynomial $J^r P(z)$ at the point h.) The point $z \in \mathbb{C}^3$ is flecnodal if $J^3 P(z)$ vanishes on a line through 0.

We can generalize the definition of flecnode in various ways. A point $z \in Z(P)$ is called r-flecnodal if there is a line through z where P vanishes to order r. In terms of jets, z is r-flecnodal for P if $J^r P(z)$ vanishes on a line through 0. A point z is doubly r-flecnodal if $J^r P(z)$ vanishes on two distinct lines through 0.

If z lies in two distinct lines in $Z(P)$, then z is clearly doubly r-flecnodal for any r. This is our special feature of points in $P_2(\mathfrak{L})$. If P vanishes on the lines of $\mathfrak{L}$, then every point of $P_2(\mathfrak{L})$ is doubly r-flecnodal for every r.

Next we want to understand if this condition is contagious. For instance, if $Z(P)$ has many doubly 10-flecnodal points, does it follow that every point of $Z(P)$ is doubly 10-flecnodal? Are there polynomials that detect doubly 10-flecnodal points in the way that SP detects flat/critical points?

In the 1800's, Salmon introduced the flecnode polynomial. He proved that for any polynomial $P \in \operatorname{Poly}(\mathbb{C}^3)$, there is another polynomial $\operatorname{Flec} P \in \operatorname{Poly}(\mathbb{C}^3)$ so that a point $z \in Z(P)$ is flecnodal if and only if $\operatorname{Flec} P(z) = 0$. He also proved that $\operatorname{Deg} \operatorname{Flec} P \le 11 \operatorname{Deg} P$. We will prove a generalization of this result for r-flecnodal points for any r, as well as a version of the result for doubly r-flecnodal points.

Let us give the generalization of Salmon's theorem for doubly r-flecnodal points. Informally, it says the following. For any polynomial $P \in \operatorname{Poly}(\mathbb{C}^3)$, there is a finite list of other polynomials called $\operatorname{Flec}_{2,r,j} P$ with $\operatorname{Deg} \operatorname{Flec}_{2,r,j} P \lesssim \operatorname{Deg} P$, and these polynomials encode which points $z \in \mathbb{C}^3$ are doubly r-frecnodal for P. A little more

precisely, for any point $z \in \mathbb{C}^3$, if you tell me, for each j, whether $\mathrm{Flec}_{2,r,j} P(z) = 0$, then I will have enough information to know whether z is doubly r-flecnodal.

Here is a formal statement of the generalized Salmon's theorem for doubly r-flecnodal points. We first define a function $v : \mathbb{C} \to \{0, 1\}$ by setting $v(0) = 0$ and $v(z) = 1$ if $z \neq 0$.

PROPOSITION 13.3. For each $r > 0$, there is an integer $J(r)$ and a subset $B_r \subset \{0, 1\}^{J(r)}$ so that the following holds. For each $P \in \mathrm{Poly}(\mathbb{C}^3)$, there is a list of other polynomials $\mathrm{Flec}_{2,r,j} P \in \mathrm{Poly}(\mathbb{C}^3)$ with $j = 1, ..., J(r)$, so that z is doubly r-flecnodal if and only if

$$\big(v(\mathrm{Flec}_{2,r,1} P(z)), ..., v(\mathrm{Flec}_{2,r,J(r)} P(z))\big) \in B_r \subset \{0, 1\}^{J(r)}.$$

Moreover, for each j,

$$\mathrm{Deg}(\mathrm{Flec}_{2,r,j} P) \leq C(r) \,\mathrm{Deg}\, P.$$

This Proposition is good enough to show that being doubly r-flecnodal is quite contagious. In a sense that we will make precise below, if $Z(P)$ has too many points that are doubly 10-flecnodal, then almost every point of $Z(P)$ is doubly 10-flecnodal.

The proof of Proposition 13.3 is based on projection theory. To see the basic issue, we start with flecnodal points. Recall that a point z is flecnodal for P if $J^3 P(z) \in \mathrm{Poly}_3(\mathbb{C}^3)$ vanishes on a line through 0. Let $\mathrm{Flec} \subset \mathrm{Poly}_3(\mathbb{C}^3)$ be the set of polynomials that vanish on a line through 0. A point z is flecnodal if and only if $J^3 P(z) \in \mathrm{Flec}$. Recall that $\mathbb{CP}^2$ is the set of lines through 0 in $\mathbb{C}^3$. When we talk about flecnodal points, the following set naturally appears:

$$V := \{(P, l) \in \mathrm{Poly}_3(\mathbb{C}^3) \times \mathbb{CP}^2 | P \text{ vanishes on } l\}.$$

It is not hard to show that this set V is an algebraic subset of $\mathbb{C}^3 \times \mathbb{CP}^2$. But what about the set Flec? The set Flec is the projection of V to $\mathbb{C}^3$. Is Flec also an algebraic set?

Projection theory studies the structure of this type of set. Here is one of the basic questions of the theory. Given an algebraic set $Y \subset \mathbb{C}^m \times \mathbb{C}^n$, if we let $\pi(Y)$ be the projection of Y to $\mathbb{C}^m$, what kind of set is $\pi(Y)$? In general $\pi(Y)$ is not an algebraic set, but a fundamental theorem says that $\pi(Y)$ is defined in terms of finitely many equations and non-equations. This is Chevalley's projection theorem, the main tool from algebraic geometry that we use in the chapter. It is a flexible tool that enables one to prove a variety of results in the flavor of Proposition 13.3. Chevalley's projection theorem works over $\mathbb{C}$ but not over $\mathbb{R}$ – this is the reason that we work over $\mathbb{C}$ throughout this chapter.

We now return to the outline of the proof of Theorem 13.33. We know that the lines $\mathfrak{L}$ lie in a surface $Z(P)$ of small degree and that each point of $P_2(\mathfrak{L})$ is doubly r-flecnodal. Any value of $r \geq 3$ will work in the rest of the argument, so we now focus on $r = 3$. A doubly 3-flecnodal point is just called doubly flecnodal. With Proposition 13.3 in hand, it is not hard to prove that being doubly flecnodal is contagious. If $P_2(\mathfrak{L})$ is too big, we will be able to prove that (almost) every point of $Z(P)$ is doubly flecnodal.

To finish the proof of Theorem 13.33, we have to classify irreducible algebraic surfaces where (almost) every point is doubly flecnodal – proving that every such surface has degree at most 2. This requires another idea from ruled surface theory.

We recall that an algebraic surface $Z(P) \subset \mathbb{C}^3$ is called a ruled surface if each point in $Z(P)$ lies in a line in $Z(P)$. Clearly, if $Z(P)$ is ruled then every point of $Z(P)$ is flecnodal. Remarkably, the converse is also true. It was proven by Cayley and Salmon, and earlier by Monge. See [**Ko**] for more information about the history.

THEOREM 13.4. (Monge-Cayley-Salmon) If $P \in \text{Poly}(\mathbb{C}^3)$ and every point of $Z(P)$ is flecnodal, then $Z(P)$ is ruled.

This theorem may be a little surprising at first. It may easily happen that a line is tangent to a surface $Z(P)$ to order 3 but does not lie in $Z(P)$. So a point of $Z(P)$ can be flecnodal even if $Z(P)$ does not contain any lines. But if every point of $Z(P)$ is flecnodal, then the theorem says that every point of $Z(P)$ lies in a line of $Z(P)$.

This theorem has a local-to-global flavor. The hypothesis that each point of $Z(P)$ is flecnodal gives local information about $Z(P)$. We need to turn this local information into the global conclusion that $Z(P)$ contains many lines. The argument uses some simple differential geometry. We have already mentioned a simple local-to-global argument in Chapter 11, when we noted that if $Z(P) \subset \mathbb{R}^3$ is flat at each regular point (and if $Z(P)$ contains a regular point) then $Z(P)$ is a plane (see Lemma 11.13). The theorem of Monge-Cayley-Salmon has the same flavor, but it is subtler.

To prove Theorem 13.1, we need a doubly-ruled analogue of the Monge-Cayley-Salmon theorem.

PROPOSITION 13.5. If $P \in \text{Poly}(\mathbb{C}^3)$ is irreducible, and every point of $Z(P)$ is doubly flecnodal, then there is an open set $O \subset Z(P)$ so that every point of O is regular and every point of O lies in two different lines in $Z(P)$.

(This Proposition is actually easier than Theorem 13.4. We will prove Proposition 13.5, but not Theorem 13.4. For a discussion of the proof of Theorem 13.4, see [**Ko**] or [**Ka**].)

Once we have found all these lines in $Z(P)$, we will use the theory of reguli to prove that $Z(P)$ is a union of planes and reguli. (This last argument is essentially equivalent to the classification of doubly ruled surfaces mentioned above, and we will prove that result at the same time.)

Here is an outline of the Chapter. First, we will introduce projection theory and prove Chevalley's theorem. Then we will use this important tool to study doubly r-flecnodal points. After that, we will review some differential geometry and use it to prove Proposition 13.5. With these two tools available, we will be able to give a short proof of Theorem 13.33.

13.1. Projection theory

Let $\mathbb{F}$ be a field. Recall that an algebraic set in $\mathbb{F}^n$ is just the zero set of a finite list of polynomials. Suppose that Z is an algebraic set in $\mathbb{F}^m \times \mathbb{F}^n$, and we consider the projection of Z onto the second factor. Is the projection also an algebraic set?

In general the answer is no. We consider two examples. We begin working over the field $\mathbb{R}$ where everything is as simple as possible to visualize.

EXAMPLE 13.6. (Circle example) Let Z be the zero set of $x^2 + y^2 - 1$ in $\mathbb{R}^2$. If we project Z to the x-axis we get the closed segment $[-1, 1]$. This is not an algebraic set.

EXAMPLE 13.7. (Hyperbola example) Let Z be the zero set of $xy = 1$ in $\mathbb{R}^2$. If we project Z to the x-axis, we get $\mathbb{R} \setminus \{0\}$. This is not an algebraic set.

What would happen if we work over $\mathbb{C}$ instead of $\mathbb{R}$? The example with the circle gets better. If we let Z be the zero set of x^2+y^2-1 in $\mathbb{C}^2$, then the projection of Z to the x axis is $\mathbb{C}$. But the hyperbola example is the same as before – if we work over $\mathbb{C}$, the image of the projection is $\mathbb{C} \setminus \{0\}$. The set $\mathbb{C} \setminus \{0\}$ is not an algebraic set, but it is an example of a slightly more general object called a constructible set.

A constructible set is determined by the vanishing or non-vanishing of finitely many polynomials. More precisely, this means the following. Let $\mathbb{F}$ be a field. As above, define $v : \mathbb{F} \to \{0, 1\}$ by setting $v(0) = 0$ and $v(z) = 1$ for $z \neq 0$ A constructible set $Y \subset \mathbb{F}^n$ is described by the following data:

- A list of polynomials $P_j \in \mathbb{F}[z_1, ..., z_n]$, for $j = 1, ..., J$.
- A subset $B \subset \{0, 1\}^J$.

The corresponding constructible set Y is defined as

$$Y := \{z \in \mathbb{F}^n | (v(P_1(z)), ..., v(P_J(z))) \in B\}.$$

We denote this constructible set as $Y(P_1, ..., P_J; B)$.

For instance, the set $\mathbb{C} \setminus \{0\}$ is not an algebraic set but it is a constructible set. Chevalley proved that over $\mathbb{C}$, any projection of an algebraic set is constructible. More generally, any projection of a constructible set in constructible. This theorem is one of the most fundamental results in projection theory.

THEOREM 13.8. (Chevalley) Suppose that Y is a constructible set in $\mathbb{C}^m \times \mathbb{C}^n$. Let $\pi : \mathbb{C}^m \times \mathbb{C}^n \to \mathbb{C}^m$ be the projection to the first factor. Then $\pi(Y) \subset \mathbb{C}^m$ is a constructible set.

We begin with some simple facts about constructible sets over any field $\mathbb{F}$.

LEMMA 13.9. For any field $\mathbb{F}$, constructible sets in $\mathbb{F}^n$ enjoy the following properties.

(1) The complement of a constructible set is constructible.
(2) A finite union or intersection of constructible sets is constructible.
(3) If $f : \mathbb{F}^m \to \mathbb{F}^n$ is a polynomial map and $Y \subset \mathbb{F}^n$ is constructible, then $f^{-1}(Y)$ is constructible.

PROOF. Suppose that $Y = Y(P_1, ..., P_J; B)$ is a constructible set, where $P_j \in \text{Poly}(\mathbb{F}^n)$ and $B \subset \{0, 1\}^J$. Then the complement of Y is $Y(P_1, ..., P_J; B^c)$, which is a constructible set.

Next, suppose that Y_a is a constructible set defined using the polynomials $P_{j,a}$, for $a = 1, ..., A$. In other words, membership in Y_a depends only on the values of $v(P_{j,a})$. Then membership in the union of Y_a depends only on the values of $v(P_{j,a})$. Therefore, $\cup_{a=1}^A Y_a$ is a constructible set. By the same argument, the intersection of Y_a is a constructible set.

Finally, suppose that $Y(P_1, ..., P_J; B)$ is a constructible set in $\mathbb{F}^n$, and suppose that $f : \mathbb{F}^m \to \mathbb{F}^n$ is a polynomial map. Then $P_j \circ f$ are polynomials on $\mathbb{F}^m$, and $f^{-1}(Y) = Y(P_1 \circ f, ..., P_j \circ f; B)$ is a constructible set in $\mathbb{F}^m$.

□

The definition of constructible set makes sense in any vector space. In particular, $\text{Poly}_d(\mathbb{C}^n)$ is a complex vector space, and it makes sense to talk about

constructible subsets of $\text{Poly}_d(\mathbb{C}^n)$. This will come up during the proof of Theorem 13.8 and also later in the chapter.

Using induction on the dimension, Theorem 13.8 follows rather quickly from a certain 1-dimensional result. Consider the following condition on some polynomials P_i and Q_j:

(13.2) There exists $z \in \mathbb{C}$ so that $P_i(z) = 0$ for all i and $Q_j(z) \neq 0$ for all j.

We will be interested in the set of polynomials that obey this condition, and with some bounds on their degrees. Given degrees $d_1, d_2, ..., d_I$, and $e_1, e_2, ..., e_J$, we define $Y(\vec{d}, \vec{e})$ to be the set of tuples $(P_1, ...P_I, Q_1, ..., Q_J)$ where $P_i \in \text{Poly}_{d_i}(\mathbb{C})$ and $Q_j \in \text{Poly}_{e_j}(\mathbb{C})$ obeying Condition 13.2. In symbols, $Y(\vec{d}, \vec{e})$ is

$$\left\{ (P_1, ..., P_I, Q_1, ..., Q_J) \in \prod_{i=1}^{I} \text{Poly}_{d_i}(\mathbb{C}) \times \prod_{j=1}^{J} \text{Poly}_{e_j}(\mathbb{C}) \text{ obeying Condition 13.2} \right\}.$$

The main step in the proof of Chevalley's theorem is to prove that $Y(\vec{d}, \vec{e})$ is a constructible subset of $\prod_{i=1}^{I} \text{Poly}_{d_i}(\mathbb{C}) \times \prod_{j=1}^{J} \text{Poly}_{e_j}(\mathbb{C})$. We state this as a proposition.

PROPOSITION 13.10. For any degrees $d_1, ..., d_I$, $e_1, ..., e_J$, the set $Y(\vec{d}, \vec{e})$ is a constructible subset of $\prod_{i=1}^{I} \text{Poly}_{d_i}(\mathbb{C}) \times \prod_{j=1}^{J} \text{Poly}_{e_j}(\mathbb{C})$.

PROOF OF THEOREM 13.8 USING PROPOSITION 13.10. It suffices to prove that the projections $\pi : \mathbb{C}^n \to \mathbb{C}^{n-1}$ send constructible sets to constructible sets. Using this result repeatedly, we can handle projections from $\mathbb{C}^n$ to $\mathbb{C}^{n-k}$ for any k, which gives the general theorem. We let $\pi : \mathbb{C}^n \to \mathbb{C}^{n-1}$ be the projection $\pi(z_1, ..., z_{n-1}, z_n) = (z_1, ..., z_{n-1})$.

Suppose that $Y \subset \mathbb{C}^n$ is constructible. By definition, $Y = \{z \in \mathbb{C}^n | v(P_j) \in B\}$ for some polynomials $P_1, ..., P_J$ and some subset $B \subset \{0,1\}^n$. The set Y is a finite union $Y = \cup_{b \in B} Y_b$, where $Y_b = \{z \in \mathbb{C}^n | v(P_j) = b \in \{0,1\}^n\}$. It suffices to prove that the projection of Y_b is constructible, because $\pi(Y) = \cup_{b \in B} \pi(Y_b)$, and a finite union of constructible sets is constructible.

After relabeling the polynomials P_j, we can write each set Y_b in the following form

$$\{z \in \mathbb{C}^n | P_i(z) = 0 \text{ for } 1 \leq i \leq I \text{ and } Q_j(z) \neq 0 \text{ for } 1 \leq j \leq J\}$$

We consider each polynomial $P_i(z)$ as a polynomial in the last coordinate z_n depending upon the other coordinates:

$$P_i(z) = \sum_{k=0}^{d_i} a_{i,k}(z_1, ..., z_{n-1}) z_n^k = P_{i, z_1, ..., z_{n-1}}(z_n).$$

In this formula the coefficients $a_{i,k}(z_1, ..., z_{n-1})$ are polynomials in $z_1, ..., z_{n-1}$. Similarly, we write $Q_j(z)$ as

$$Q_j(z) = \sum_{k=0}^{e_j} b_{j,k}(z_1, ..., z_{n-1}) z_n^k = Q_{j, z_1, ..., z_{n-1}}(z_n).$$

The point $(z_1, ..., z_{n-1})$ lies in $\pi(Y_b)$ if and only if there exists a $z_n \in \mathbb{C}$ so that $P_i(z_1, ..., z_{n-1}, z_n) = 0$ and $Q_j(z_1, ..., z_{n-1}, z_n) \neq 0$. By Proposition 13.10, this occurs if and only if the polynomials $P_{i, z_1, ..., z_{n-1}}$ and $Q_{j, z_1, ..., z_{n-1}}$ lie in $Y(\vec{d}, \vec{e})$, a

constructible set. Now the map f sending $(z_1, ..., z_{n-1})$ to the list of polynomials $P_{i,z_1,...,z_{n-1}}, Q_{j,z_1,...,z_{n-1}}$ in $\prod_{i=1}^{I} \mathrm{Poly}_{d_i}(\mathbb{C}) \times \prod_{j=1}^{J} \mathrm{Poly}_{e_j}(\mathbb{C})$ is a polynomial map. Therefore, $\pi(Y_b) = f^{-1}\left(Y(\vec{d}, \vec{e})\right)$ is a constructible set.

□

Now we turn to the proof of Proposition 13.10:

PROOF. We begin with a lemma.

LEMMA 13.11. Let $M_{m,n}(\mathbb{C})$ denote the vector space of $m \times n$ matrices with entries in $\mathbb{C}$. The subset of matrices of rank r is a constructible set.

PROOF. A matrix in $M_{m,n}(\mathbb{C})$ has rank r if and only if the determinant of every $(r+1) \times (r+1)$ minor vanishes and the determinant of some $r \times r$ minor does not vanish. Each determinant of a minor of the matrix is a polynomial on the vector space $M_{m,n}(\mathbb{C})$, and being rank r is a Boolean condition depending on the vanishing or non-vanishing of finitely many polynomials. □

We will apply this lemma to the matrix of a linear map that describes multiplication of polynomials. Given polynomials $P_1, ..., P_I \in \mathrm{Poly}(\mathbb{C})$ of degrees $d_1, ..., d_I$, and given a degree $e \geq \max d_i$, we define a linear map

$$M[P_1, ..., P_I, e] : \prod_{i=1}^{I} \mathrm{Poly}_{e-d_i}(\mathbb{C}) \to \mathrm{Poly}_e(\mathbb{C}),$$

by the formula

$$M[P_1, ..., P_I, e](g_1, ..., g_I) := P_1 g_1 + ... + P_I g_I.$$

For large enough e, the rank of $M[P_1, ..., P_I, e]$ is related to the greatest common divisor of $P_1, ..., P_I$, written $\gcd(P_1, ..., P_I)$.

LEMMA 13.12. Suppose that $P_i \in \mathrm{Poly}_{d_i}(\mathbb{C})$ with $\mathrm{Deg}\, P_i = d_i$. If $e \geq 2 \max_i d_i$, then the image of $M[P_1, ..., P_I, e]$ is exactly the multiples of $\gcd(P_1, ..., P_I)$ in $\mathrm{Poly}_e(\mathbb{C})$. In particular, if $e \geq 2 \max_i d_i$, then $\mathrm{Rank}\, M[P_1, ..., P_I, e] = e + 1 - \mathrm{Deg}(\gcd(P_1, ..., P_I))$.

PROOF. Since $M[P_1, ..., P_I, e](g_1, ..., g_I) := P_1 g_1 + ... + P_I g_I$, the image of $M[P_1, ..., P_I, e]$ is contained in the multiples of $\gcd(P_1, ..., P_I)$. It remains to check that the image of $M[P_1, ..., P_I, e]$ contains the multiples of $\gcd(P_1, ..., P_I)$ in $\mathrm{Poly}_e(\mathbb{C})$.

We do the proof by induction on I. We begin with $I = 2$, which we use as a base case. The dimension of the domain of $M[P_1, P_2, e]$ is $\mathrm{Dim}\, \mathrm{Poly}_{e-d_1}(\mathbb{C}) + \mathrm{Dim}\, \mathrm{Poly}_{e-d_2}(\mathbb{C}) = 2e - d_1 - d_2 + 2$.

$$\mathrm{Dim}(\mathrm{Domain}\, M[P_1, P_2, e]) = 2e - d_1 - d_2 + 2. \tag{13.3}$$

Next we determine the kernel of $M[P_1, P_2, e]$. We let

$$P := \gcd(P_1, P_2).$$

$$d := \mathrm{Deg}\, P.$$

Therefore, we have

$$P_1 = P \cdot \tilde{P}_1; P_2 = P \cdot \tilde{P}_2.$$

$$\mathrm{Deg}\, \tilde{P}_i = d_i - d.$$

Now we define a linear map

$$A : \mathrm{Poly}_k(\mathbb{C}) \to \mathrm{Poly}_{e-d_1}(\mathbb{C}) \times \mathrm{Poly}_{e-d_2}(\mathbb{C}),$$

where

$$A(h) = (\tilde{P}_2 h, -\tilde{P}_1 h),$$

and

$$k = e - d_1 - d_2 + d.$$

(The condition that $e \geq 2 \max_i d_i$ guarantees that $k \geq 0$.)

We claim that A is an isomorphism from $\mathrm{Poly}_k(\mathbb{C})$ to $\mathrm{Ker}\, M[P_1, P_2, e]$.

First we check that A does indeed map $\mathrm{Poly}_k(\mathbb{C})$ into $\mathrm{Poly}_{e-d_1}(\mathbb{C}) \times \mathrm{Poly}_{e-d_2}(\mathbb{C})$. We have $\mathrm{Deg}\, \tilde{P}_2 h = \mathrm{Deg}\, \tilde{P}_2 + \mathrm{Deg}\, h \leq (d_2 - d) + (e - d_1 - d_2 + d) = e - d_1$. Similarly, $\mathrm{Deg}\, \tilde{P}_1 h \leq (d_1 - d) + (e - d_1 - d_2 + d) = e - d_2$.

Next we check that the image of A lies in $\mathrm{Ker}\, M[P_1, P_2, e]$:

$$\begin{aligned} M[P_1, P_2, e](A(h)) &= M[P_1, P_2, e](\tilde{P}_2 h, -\tilde{P}_1 h) \\ &= P_1 \tilde{P}_2 h - P_2 \tilde{P}_1 h = P \tilde{P}_1 \tilde{P}_2 h - P \tilde{P}_1 \tilde{P}_2 h = 0. \end{aligned}$$

Next we check that the image of A contains $\mathrm{Ker}\, M[P_1, P_2, e]$. Suppose that $(g_1, g_2) \in \mathrm{Ker}\, M[P_1, P_2, e]$. We have

$$0 = P_1 g_1 + P_2 g_2 = P(\tilde{P}_1 g_1 + \tilde{P}_2 g_2).$$

Since $\gcd(\tilde{P}_1, \tilde{P}_2) = 1$, it follows that $\tilde{P}_2$ divides g_1 and $\tilde{P}_1$ divides g_2. So we can write $g_1 = \tilde{P}_2 h_1$ and $g_2 = \tilde{P}_1 h_2$. Now the equation $\tilde{P}_1 g_1 + \tilde{P}_2 g_2 = 0$ becomes $\tilde{P}_1 \tilde{P}_2 (h_1 + h_2) = 0$, and so $h_2 = -h_1$. We define h to be h_1, and so $(g_1, g_2) = (\tilde{P}_2 h, -\tilde{P}_1 h)$.

To see that (g_1, g_2) is in the image of A, it just remains to check that $\mathrm{Deg}\, h \leq k$. We note

$$\mathrm{Deg}\, h + \mathrm{Deg}\, \tilde{P}_2 = \mathrm{Deg}\, g_1 \leq e - d_1,$$

and so $\mathrm{Deg}\, h + (d_2 - d) \leq e - d_1$, which gives

$$\mathrm{Deg}\, h \leq e - d_1 - d_2 + d = k.$$

We have now proven that the image of A is exactly $\mathrm{Ker}\, M[P_1, P_2, e]$. The map A is clearly injective, and so A is an isomorphism from $\mathrm{Poly}_k(\mathbb{C})$ to $\mathrm{Ker}\, M[P_1, P_2, e]$. In particular, we see that

$$\mathrm{Dim}\, \mathrm{Ker}\, M[P_1, P_2, e] = \mathrm{Dim}\, \mathrm{Poly}_k(\mathbb{C}) = k + 1 = e - d_1 - d_2 + d + 1.$$

Comparing the dimension of the domain and the dimension of the kernel, we see that the dimension of the range of $M[P_1, P_2, e]$ is

$$\mathrm{Dim}(\mathrm{Range}\, M[P_1, P_2, e]) = e - d + 1. \tag{13.4}$$

We already know that $\mathrm{Range}\, M[P_1, P_2, e]$ is contained in the multiples of $P = \gcd(P_1, P_2)$ in $\mathrm{Poly}_e(\mathbb{C})$. But the dimension of this space of multiples is $e - d + 1$. Therefore, $\mathrm{Range}\, M[P_1, P_2, e]$ is exactly the multiples of $\gcd(P_1, P_2)$ in $\mathrm{Poly}_e(\mathbb{C})$.

This proves our result for $I = 2$. Now we prove the result for all I by induction. Suppose that $I \geq 3$ and $e \geq 2 \max d_i$. We want to understand the image of $M[P_1, ..., P_I, e]$, that is the set of polynomials of the form

$$P_1 g_1 + ... + P_{I-1} g_{I-1} + P_I g_I, \mathrm{Deg}\, g_i \leq e - d_i.$$

We start by considering the possible values of the sum of the last two terms: $P_{I-1} g_{I-1} + P_I g_I$. By the case $I = 2$, we know that this sum could be any multiple

of $\gcd(P_{I-1}, P_I)$ in $\text{Poly}_e(\mathbb{C})$. In other words, this last sum could be any polynomial of the form $\gcd(P_{I-1}, P_I)\bar{g}$ where $\text{Deg}\,\bar{g} \leq e - \text{Deg}(\gcd(P_1, P_2))$. But this means that the range is the set of polynomials of the form:

$$P_1 g_1 + ... + P_{I-2} g_{I-2} + \gcd(P_{I-1}, P_I)\bar{g}, \text{Deg}\, g_i \leq e - d_i, \text{Deg}\,\bar{g} \leq e - d.$$

In other words,

$$\text{Range}\, M[P_1, ..., P_I, e] = \text{Range}\, M[P_1, ..., P_{I-2}, \gcd(P_{I-1}, P_I), e]. \tag{13.5}$$

By induction on I, we know that this image is the set of multiples of
$\gcd(P_1, ..., P_{I-2}, \gcd(P_{I-1}, P_I)) = \gcd(P_1, ..., P_I)$.
This proves the lemma.

□

Combining the last two lemmas, we get the following corollary:

COROLLARY 13.13. For any degrees $d_1, ..., d_I, d$, the subset of tuples $(P_1, ..., P_I) \in \text{Poly}_{d_1}(\mathbb{C}) \times ... \times \text{Poly}_{d_I}(\mathbb{C})$ so that $\text{Deg}\, P_i = d_i$ for all i and $\text{Deg}(\gcd(P_1, ..., P_I)) = d$ is a constructible set.

PROOF. Pick $e \geq 2 \max d_i$. Suppose that $\text{Deg}\, P_i = d_i$ for all i. By Lemma 13.12, $\text{Deg}(\gcd(P_1, ..., P_I)) = d$ if and only if $\text{Rank}\, M[P_1, ..., P_I, e] = e - d + 1$. We are going to apply Lemma 13.11, which says that the set of $m \times n$ matrices with rank r is a constructible subset of $M_{m,n}(\mathbb{C})$. In order to do this, we have to write the linear operator $M[P_1, ..., P_I, e]$ as a matrix.

To write $M[P_1, ..., P_I, e]$ as a matrix, we have to choose a basis for the domain and target. Recall that the domain is the direct sum of $\text{Poly}_{e-d_i}(\mathbb{C})$, and the target is $\text{Poly}_e(\mathbb{C})$. The space $\text{Poly}_d(\mathbb{C})$ has a natural basis of monomials: $1, z, z^2, ...$ We use this monomial basis for $\text{Poly}_e(\mathbb{C})$ and for each $\text{Poly}_{e-d_i}(\mathbb{C})$, giving a basis for the domain and target. We define $\mu(P_1, ..., P_I)$ to be the matrix for $M[P_1, ..., P_I, e]$ written in this basis. Each entry of $\mu(P_1, ..., P_I)$ is just a coefficient of one of the polynomials P_i. The entry in the column of the matrix corresponding to the z^j term of $g_i \in \text{Poly}_{e-d_i}(\mathbb{C})$ and the row of the matrix corresponding to the z^k term of the output is the coefficient of z^{k-j} in P_i. Therefore, $\mu : \prod_i \text{Poly}_{d_i}(\mathbb{C}) \to M_{m,n}(\mathbb{C})$ is a linear map. (Here m and n are the dimensions of the matrix $\mu(P_1, ..., P_I)$: $m = \text{Dim}\,\text{Poly}_e(\mathbb{C})$ and $n = \sum_i \text{Dim}\,\text{Poly}_{e-d_i}(\mathbb{C})$.)

By Lemma 13.11, the set of matrices of rank $e - d + 1$ is a constructible subset of $M_{m,n}(\mathbb{C})$. Since μ is linear, the set of $(P_1, ..., P_I)$ so that $\text{Rank}\, M[P_1, ..., P_I, e] = e - d + 1$ is a constructible subset of $\prod_{i=1}^I \text{Poly}_{d_i}(\mathbb{C})$. The set of $(P_1, ..., P_I)$ with $\text{Deg}\, P_i = d_i$ is clearly constructible. The set of $(P_1, ..., P_I)$ so that $\text{Deg}\, P_i = d_i$ and $\text{Deg}\gcd(P_1, ..., P_I) = d$ is the intersection of these two constructible sets, which is constructible. □

COROLLARY 13.14. For any degrees $d_1, ..., d_I, d$, the subset of tuples $(P_1, ..., P_I) \in \text{Poly}_{d_1}(\mathbb{C}) \times ... \times \text{Poly}_{d_I}(\mathbb{C})$ so that $\text{Deg}(\gcd(P_1, ..., P_I)) = d$ is a constructible set.

PROOF. We let $\text{Poly}_{=d}(\mathbb{C})$ be the set of polynomials of degree exactly d. The set $\text{Poly}_{=d}(\mathbb{C})$ is not a vector space (for instance, it does not contain zero). A constant polynomial has degree 0. We make the convention that the zero polynomial has degree -1. Then we can decompose the space $\text{Poly}_d(\mathbb{C})$ by degree as follows:

$$\text{Poly}_d(\mathbb{C}) = \cup_{e=-1}^d \text{Poly}_{=e}(\mathbb{C}).$$

It's easy to see that each subset $\text{Poly}_{=e}(\mathbb{C}) \subset \text{Poly}_d(\mathbb{C})$ is constructible.

We can decompose $\prod_i \mathrm{Poly}_{d_i}(\mathbb{C})$ as the disjoint union of $\prod_i \mathrm{Poly}_{=e_i}(\mathbb{C})$ for some $-1 \leq e_i \leq d_i$.

It suffices to prove that the set of $(P_1, ..., P_I) \in \prod_i \mathrm{Poly}_{=e_i}(\mathbb{C})$ with $\mathrm{Deg}\,\gcd(P_1, ..., P_I) = d$ is a constructible set. But this follows from Corollary 13.13. □

At this point it is convenient to define a constructible function. We say that a function F on $\mathbb{C}^n$ is constructible if it takes finitely many values, and the preimage of each value is a constructible set. We have just proved that $\mathrm{Deg}(\gcd(P_1, ..., P_I))$ is a constructible function on $\mathrm{Poly}_{d_1}(\mathbb{C}) \times ... \times \mathrm{Poly}_{d_I}(\mathbb{C})$. Because finite unions and intersections of constructible sets are constructible, finite sums and products of constructible functions are constructible.

Recall that we write $Z(P_1, ..., P_I)$ for the set of $z \in \mathbb{C}$ where $P_1(z) = ... = P_I(z) = 0$. Next, we will prove that the cardinality $|Z(P_1, ..., P_I)|$ is a constructible function on $\mathrm{Poly}_{d_1}(\mathbb{C}) \times ... \times \mathrm{Poly}_{d_I}(\mathbb{C})$. Up to this point in the argument, we have not used any special properties of $\mathbb{C}$: Corollary 13.14 is true over any field with the same proof. Now we will use that the complex numbers are algebraically closed. To see how this is relevant, we observe that $Z(P_1, ..., P_I)$ is non-empty in $\mathbb{C}$ if and only if $\mathrm{Deg}\,\gcd(P_1, ..., P_I) > 0$. On the one hand, if $\beta \in Z(P_1, ..., P_I)$, then $z - \beta$ divides $P_1, ..., P_I$, and so $\mathrm{Deg}\,\gcd(P_1, ..., P_I) > 0$. On the other hand, suppose that $P = \gcd(P_1, ..., P_I)$ has degree > 0. Because $\mathbb{C}$ is algebraically closed, P factors as $c\prod_k (z - \beta_k)^{\mu_k}$, where $c \neq 0$, $\beta_k \in \mathbb{C}$, and $\mu_k \geq 1$. In particular, $\beta_1 \in Z(P_1, ..., P_I)$. This shows that the set of $(P_1, ..., P_I) \in \mathrm{Poly}_{d_1}(\mathbb{C}) \times ... \times \mathrm{Poly}_{d_I}(\mathbb{C})$ so that $Z(P_1, ..., P_I)$ is non-empty is a constructible set. With a little bit more work, we will prove that $|Z(P_1, ..., P_I)|$ is a constructible function.

LEMMA 13.15. For any degrees $d_1, ..., d_I$, the function $|Z(P_1, ..., P_I)|$ is a constructible function on $\mathrm{Poly}_{d_1}(\mathbb{C}) \times ... \times \mathrm{Poly}_{d_I}(\mathbb{C})$.

PROOF. Suppose that $\gcd(P_1, ..., P_I) = \prod_{k=1}^K (z - \beta_k)^{\mu_k}$, where $\mu_k \geq 1$. The degree of $\gcd(P_1, ..., P_I) = \sum_{k=1}^K \mu_k$, and $|Z(P_1, ..., P_I)| = K$.

As usual, we write P_i' for the derivative of P_i. The key observation is that

$$\gcd(P_1, P_1', P_2, P_2', ..., P_I, P_I') = \prod_{k=1}^K (z - \beta_k)^{\mu_k - 1}. \tag{13.6}$$

Given this formula, we can solve for $|Z(P_1, ..., P_I)|$ by

$$|Z(P_1, ..., P_I)| = \mathrm{Deg}\left(\gcd(P_1, ..., P_I)\right) - \mathrm{Deg}\left(\gcd(P_1, P_1', ..., P_I, P_I')\right).$$

By Corollary 13.14, $\mathrm{Deg}(\gcd(P_1, ..., P_I))$ and $\mathrm{Deg}(\gcd(P_1, P_1', ..., P_I, P_I'))$ are constructible functions, and so $|Z(P_1, ..., P_I)|$ is constructible as well.

So it only remains to check Equation 13.6 for $\gcd(P_1, P_1', ..., P_I, P_I')$. We know that $(z - \beta_k)^{\mu_k}$ divides P_i for every i, k. By the Leibniz formula, $(z - \beta_k)^{\mu_k - 1}$ divides P_i' for every i, k. Hence $\prod_k (z - \beta_k)^{\mu_k - 1}$ divides every P_i and every P_i'.

Finally, we have to check that $(z - \beta_k)^{\mu_k}$ does not divide every P_i'. Fix k. Since $\gcd(P_1, ..., P_I) = \prod_k (z - \beta_k)^{\mu_k}$, there must be some i so that $(z - \beta_k)^{\mu_k + 1}$ does not divide P_i. Then $P_i = (z - \beta_k)^{\mu_k}\tilde{P}_i$,where $(z - \beta_k)$ does not divide $\tilde{P}_i$. By the Liebniz rule,

$$P_i' = \mu_k (z - \beta_k)^{\mu_k - 1}\tilde{P}_i + (z - \beta_k)^{\mu_k}\tilde{P}_i'.$$

From this formula, we see that $(z - \beta_k)^{\mu_k}$ does not divide P_i'. This finishes the proof of Equation 13.6 and so the proof of the lemma. □

Note that $Z(P_1, ..., P_I) = \cap_{i=1}^I Z(P_i)$. For any subset $A \subset \{1, ..., I\}$, we define

$$V_A(P_1, ..., P_I) := \cap_{i \in A} Z(P_i).$$

By Lemma 13.15, we see that for every subset $A \subset \{1, ..., I\}$, $|V_A(P_1, ..., P_I)|$ is a constructible function on $\text{Poly}_{d_1}(\mathbb{C}) \times ... \times \text{Poly}_{d_I}(\mathbb{C})$. More generally, given any sets $A_0, A_1 \subset \{1, ..., I\}$ we define

(13.7)

$$V_{A_0,A_1}(P_1, ..., P_I) := \{z \in \mathbb{C} | P_i(z) = 0 \text{ for all } i \in A_0 \text{ and } P_i(z) \neq 0 \text{ for all } i \in A_1\}.$$

LEMMA 13.16. For any degrees $d_1, ..., d_I$ and any sets $A_0, A_1 \subset \{1, ..., I\}$, the function $|V_{A_0,A_1}(P_1, ..., P_I)|$ is a constructible function on $\text{Poly}_{d_1}(\mathbb{C}) \times ... \times \text{Poly}_{d_I}(\mathbb{C})$.

PROOF. If A_0 and A_1 intersect, then $V_{A_0,A_1}(P_1, ..., P_I)$ is empty, and so the conclusion is trivial. We suppose that A_0 and A_1 are disjoint. Now we can express $|V_{A_0,A_1}(P_1, ..., P_I)|$ using the inclusion/exclusion principle:

$$|V_{A_0,A_1}| = |V_{A_0}| - \sum_{i \in A_1} |V_{A_0 \cup i}| + \sum_{i_1 \neq i_2 \in A_1} |V_{A_0 \cup i_1 \cup i_2}| - ...$$

Each term on the right-hand side is a constructible function by Lemma 13.15. Since finite sums and/or differences of constructible functions are constructible, $|V_{A_0,A_1}(P_1, ..., P_I)|$ is a constructible function. □

In particular, for any A_0 or A_1, we see that the set of $(P_1, ..., P_I) \in \text{Poly}_{d_1}(\mathbb{C}) \times ... \times \text{Poly}_{d_I}(\mathbb{C})$ so that $|V_{A_0,A_1}(P_1, ..., P_I)| \geq 1$ is a constructible subset. This proves Proposition 13.10.

□

We have now finished the proof of Chevalley's projection theorem, Theorem 13.8. The theorem is very general and versatile, in part because there are lots of constructible sets. For example, Theorem 13.8 easily implies the following more general result.

If $Y \subset \mathbb{C}^m \times \mathbb{C}^n$ and s is an integer ≥ 1, then we define

$$\pi^{(s)} Y := \{z \in \mathbb{C}^m | \text{ there exist } s \text{ distinct } w_k \in \mathbb{C}^n \text{ so that } (z, w_k) \in Y\}.$$

If $s = 1$, $\pi^{(1)} Y$ is just the projection of Y to the first factor.

COROLLARY 13.17. If $Y \subset \mathbb{C}^m \times \mathbb{C}^n$ is a constructible set and $s \geq 1$ is an integer, then $\pi^{(s)} Y \subset \mathbb{C}^m$ is also constructible.

PROOF. Starting with Y, we build up a variety of constructible sets. Throughout this argument, $z \in \mathbb{C}^m$ and w or w_k lie in $\mathbb{C}^n$.

For each $1 \leq k \leq s$, the following set is constructible:

$$A_k := \{(z, w_1, ..., w_s) | (z, w_k) \in Y\}.$$

Since finite intersections of constructible sets are constructible, the following set is constructible:

$$A := \cap_{k=1}^s A_k = \{(z, w_1, ..., w_s) | (z, w_k) \in Y \text{ for every } 1 \leq k \leq s\}.$$

On the other hand, for any $k_1 \neq k_2$, the following set is clearly constructible:

$$D_{k_1,k_2} := \{(z, w_1, ..., w_s) | w_{k_1} \neq w_{k_2}\}.$$

Taking the intersection of all the sets D_{k_1,k_2}, we see that the following set is constructible:

$$D := \{(z, w_1, ..., w_s) | \text{ the points } w_1, ..., w_s \text{ are distinct}\}.$$

Now the set $A \cap D$ is constructible. We can describe $A \cap D$ as follows:

$$A \cap D = \{(z, w_1, ..., w_s) | \text{ the points } w_1, ..., w_s \text{ are distinct and } (z, w_k) \in Y \text{ for all } 1 \leq k \leq s\}.$$

The set $\pi^{(s)}Y$ is the projection of $A \cap D$ onto the z-coordinate. By Theorem 13.8, $\pi^{(s)}Y$ is constructible. □

13.2. Flecnodes and double flecnodes

Let $P \in \text{Poly}(\mathbb{C}^3)$. Recall that a point x is flecnodal for P if there is a line through x so that P vanishes to third order at x along l. We can state this in terms of the degree 3 jet $J^3P(x)$. The point x is flecnodal if $J^3P(x)$ vanishes on a line through 0.

We say that a point x is doubly flecnodal for P if there are two different lines through x so that P vanishes to third order at x on each line. In other words, x is doubly flecnodal if $J^3P(x)$ vanishes on two distinct lines through 0.

More generally, we say that a point x is (s, r)-flecnodal if there are s distinct lines through x so that P vanishes to order r on each line. In other words, x is (s, r)-flecnodal if $J^rP(x)$ vanishes on s distinct lines through 0.

We define $\text{Flec}_{s,r} \subset \text{Poly}_r(\mathbb{C}^3)$ to be the set of polynomials that vanish on s distinct lines through 0. So x is (s, r)-flecnodal for P if and only if $J^rP(x) \in \text{Flec}_{s,r}$. Our first main result is that $\text{Flec}_{s,r}$ is a constructible set.

PROPOSITION 13.18. For any integers $s \geq 1$ and $r \geq 1$, the set $\text{Flec}_{s,r} \subset \text{Poly}_r(\mathbb{C}^3)$ is a constructible set.

PROOF. We begin by parametrizing lines through 0 in $\mathbb{C}^3$. For any $(a_1, a_2) \in \mathbb{C}^2$, we define l_{a_1,a_2} to be the line defined by the equations

$$z_1 = a_1 z_3; z_2 = a_2 z_3. \tag{13.8}$$

The lines l_{a_1,a_2} are all distinct. The set of lines $\{l_{a_1,a_2}\}$ are all of the lines through 0 that don't lie in the (z_1, z_2)-plane.

We consider the following set:

$$V := \{(P, a_1, a_2) \in \text{Poly}_r(\mathbb{C}^3) \times \mathbb{C}^2 | P \text{ vanishes on } l_{a_1,a_2}\}. \tag{13.9}$$

We will show that this set V is algebraic and hence constructible. In particular, we claim that a polynomial $P \in \text{Poly}_r(\mathbb{C}^3)$ vanishes on l_{a_1,a_2} if and only if

$$P(ta_1, ta_2, t) = 0 \text{ for each integer in the range } 1 \leq t \leq r + 1. \tag{13.10}$$

The points (ta_1, ta_2, t) all lie on l_{a_1,a_2}, so if P vanishes on l_{a_1,a_2}, then $P(ta_1, ta_2, t) = 0$ for all t. On the other hand, if $P(ta_1, ta_2, t) = 0$ for $t = 1, ..., r+1$, then P vanishes on $r + 1$ points of l_{a_1,a_2}. Since $P \in \text{Poly}_r(\mathbb{C}^3)$, P vanishes on l_{a_1,a_2}. Equations 13.10 are a finite list of polynomial equations in a_1, a_2, and the coefficients of P. Therefore, V is an algebraic set.

Now we consider $\pi^{(s)}V$: the set of polynomials $P \in \text{Poly}_r(\mathbb{C}^3)$ so that P vanishes on s distinct lines l_{a_1,a_2}. By Corollary 13.17, $\pi^{(s)}V$ is constructible. This

almost finishes the proof. Let π_0 denote the $z_1 z_2$-plane. We have now proven that the following set is constructible:

$$\text{Flec}_{s,r,\pi_0} := \{P \in \text{Poly}_r(\mathbb{C}^3) | P \text{ vanishes on } s \text{ distinct lines thru 0, not in } \pi_0\}.$$

There was nothing special about the (z_1, z_2)-plane. By changing coordinates, we can prove that $\text{Flec}_{s,r,\pi}$ is constructible for any complex 2-plane π thru 0. We let $\pi_1, ..., \pi_{2s+1}$ be 2-planes thru 0 in general position. We claim that

$$\text{Flec}_{s,r} = \cup_m \text{Flec}_{s,r,\pi_m} .$$

Indeed, suppose that P vanishes on s distinct lines through 0. Since the planes, π_m are in general position, each of these lines lies in at most two of them. Since there are $2s + 1$ planes, one of these planes, π_m, contains none of the s lines. Therefore, P lies in Flec_{s,r,π_m}.

Since a finite union of constructible sets is constructible, $\text{Flec}_{s,r}$ is constructible. □

Once we know that $\text{Flec}_{s,r}$ is constructible, we can easily prove Proposition 13.3. There is a more general Proposition that holds for any constructible set Y, and we now formulate it. Suppose that $Y \subset \text{Poly}_r(\mathbb{C}^3)$. We say that a polynomial P obeys the condition Y at a point z if and only if $J^r P(z) \in Y$.

LEMMA 13.19. Suppose that $Y \subset \text{Poly}_r(\mathbb{C}^3)$ is a constructible set. Then for any polynomial $P : \mathbb{C}^3 \to \mathbb{C}$, there is a finite list of polynomials $Y_j P$, $j = 1, ..., J(Y)$, and a subset $B_Y \subset \{0,1\}^{J(Y)}$ obeying the following:

- $\text{Deg } Y_j P \leq C(Y) \text{ Deg } P$.
- The polynomial P obeys condition Y at a point z if and only if

$$\big(v(Y_1 P(z)), ..., v(Y_J P(z))\big) \in B_Y.$$

PROOF. Since Y is a constructible set, there is a finite list of polynomials f_j on $\text{Poly}_r(\mathbb{C}^3)$ and a subset $B_Y \subset \{0,1\}^{J(Y)}$ so that $w \in Y$ if and only if $v(f_j(w)) \in B_Y$. The polynomial P obeys condition Y at a point z if and only if $v(f_j(J^r P(z)) \in B_Y$.

We define $Y_j P(z) = f_j(J^r T(z))$. So P obeys condition Y at z if and only if $v(Y_j P(x)) \in B_Y$.

We note that $J^r P : \mathbb{C}^3 \to \text{Poly}_r(\mathbb{C}^3)$ is a vector-valued polynomial of degree $\leq \text{Deg } P$. (Each coefficient of $J^r P$ is a constant factor times a derivative $\nabla_I P$ for some multi-index I, and each $\nabla_I P$ is a polynomial of degree $\leq \text{Deg } P$.) We let $C(Y)$ be the maximal degree of the polynomials f_j. Then $Y_j P$ is a polynomial of degree $\leq C(Y) \text{ Deg } P$. □

13.3. A definition of almost everywhere

Our next goal is to study the contagious properties of (s, r)-flecnodal points. If $Z(P)$ is a low degree surface with many (s, r)-flecnodal points, does it follow that every point of $Z(P)$ is (s, r)-flecnodal? Under appropriate conditions, we will prove that "almost every point" of $Z(P)$ is (s, r)-flecnodal. In this section, we introduce an appropriate notion of almost every point.

If $P \in \text{Poly}(\mathbb{C}^n)$ is an irreducible polynomial, we say that a condition holds at almost every point of $Z(P) \subset \mathbb{C}^n$ if the set of points $z \in Z(P)$ where the condition fails to hold is contained in $Z(Q)$ for some polynomial Q which is not divisible by P.

To illustrate the definition, we prove that almost every point of an irreducible surface is regular.

LEMMA 13.20. For any irreducible $P \in \text{Poly}(\mathbb{C}^n)$, almost every point of $Z(P)$ is regular.

PROOF. Consider the partial derivatives $\partial_i P$. Since P is irreducible, $\partial_i P$ does not divide P. We can assume that P is not constant, and so we can assume that for some i, $\partial_i P$ is not the zero polynomial. But every point in $Z(P) \setminus Z(\partial_i P)$ is regular. Therefore, almost every point of $Z(P)$ is regular. □

To show that the definition of almost every point is reasonable, we prove that if a condition holds at almost every point of $Z(P) \subset \mathbb{C}^n$, then there is a point of $Z(P)$ where it holds.

PROPOSITION 13.21. Suppose that $P \in \text{Poly}(\mathbb{C}^n)$ is irreducible, and that $Q \in \text{Poly}(\mathbb{C}^n)$ is not divisible by P. Then, there is a point in $Z(P) \setminus Z(Q)$.

PROOF. Let $I(Z(P))$ be the ideal of polynomials that vanish on $Z(P)$. By the Hilbert Nullstellensatz, $I(Z(P))$ is the radical of the ideal (P). (The reader can find a proof of the Hilbert Nullstellensatz as Theorem 1.5 in Chapter 9 of [**Lan**]. There is also a good discussion of the Nullstellensatz in Section 1.6 of [**Ei**].) We claim that since P is irreducible, $I(Z(P))$ is actually equal to (P). Suppose that $Q \in I(Z(P))$. By the definition of a radical, $Q^r \in (P)$ for some r. In other words, P divides Q^r. Since P is irreducible, and since there is unique factorization in the ring $\mathbb{C}[z_1, ..., z_n]$, P divides Q. This shows that the ideal $I(Z(P))$ is equal to (P).

Now we prove the Proposition. We suppose that Q is not divisible by P. In other words, Q is not in $(P) = I(Z(P))$. Since $Q \notin I(Z(P))$, Q does not vanish on $Z(P)$. In other words, there is a point in $Z(P) \setminus Z(Q)$. □

(We remark that this result does not hold over $\mathbb{R}$. We need $\mathbb{C}$ to be algebraically closed in order to use the Nullstellensatz.)

Combining the last two results, we get the following corollary:

COROLLARY 13.22. For any irreducible $P \in \text{Poly}(\mathbb{C}^n)$, $Z(P)$ contains a regular point.

The definition of almost every point also behaves well when we intersect two sets.

LEMMA 13.23. Suppose that $P \in \text{Poly}(\mathbb{C}^n)$ is irreducible and that $A_1, A_2 \subset Z(P)$ each contain almost every point of $Z(P)$. Then $A_1 \cap A_2$ contains almost every point of $Z(P)$.

PROOF. We know that $Z(P) \setminus A_j$ is contained in $Z(Q_j)$ where P doesn't divide Q_j. Therefore, $Z(P) \setminus (A_1 \cap A_2) = (Z(P) \setminus A_1) \cup (Z(P) \setminus A_2)$ is contained in $Z(Q_j) \cup Z(Q_2) = Z(Q_1 Q_2)$. Since there is unique factorization in the polynomial ring $\mathbb{C}[z_1, ..., z_n]$, P does not divide $Q_1 Q_2$. □

COROLLARY 13.24. Suppose that $P \in \text{Poly}(\mathbb{C}^n)$ is irreducible. Let $Y \subset \text{Poly}_r(\mathbb{C}^n)$ be any subset. Suppose that $J^r P(z) \in Y$ for almost every point $z \in Z(P)$. Then there is an open subset $O \subset Z(P)$ so that every point of O is regular and $J^r P(z) \in Y$ for every $z \in O$.

Remark. When we talk about open and closed subsets, we are using the Euclidean topology on $\mathbb{C}^n$.

PROOF. We know that $J^r P(z) \in Y$ for almost every point $z \in Z(P)$. We also know that almost every point of $Z(P)$ is regular. By Lemma 13.23, at almost every point $z \in Z(P)$, $J^r P(z) \in Y$ and z is regular. In other words, there is some polynomial Q, not divisible by P, so that for every $z \in Z(P) \setminus Z(Q)$, $J^r P(z) \in Y$ and z is a regular point of $Z(P)$.

By Proposition 13.21, there is at least one point $z_0 \in Z(P) \setminus Z(Q)$. Since Q is continuous, we can find a small radius r so that every point of $Z(P) \cap B(z_0, r)$ lies in $Z(P) \setminus Z(Q)$. These points are all regular, so we know that $Z(P) \cap B(z_0, r)$ is a complex submanifold of (complex) dimension $n-1$. We let $O := Z(P) \cap B(z_0, r)$. We know that for every $z \in O$, $J^r P(z) \in Y$. □

13.4. Constructible conditions are contagious

We are now ready to study the contagious properties of constructible conditions.

If a polynomial P obeys a constructible condition Y at too many points along a line, then it obeys Y at all but finitely many points of the line.

LEMMA 13.25. Suppose that $Y \subset \mathrm{Poly}_r(\mathbb{C}^3)$ is a constructible condition, for some $r \geq 0$. Then there is a constant $K(Y)$ so that the following holds. Suppose that $l \subset \mathbb{C}^3$ is a line. Suppose that $P : \mathbb{C}^3 \to \mathbb{C}$ is a polynomial. If P obeys condition Y at $> K(Y) \operatorname{Deg} P$ points of l, then P obeys condition Y at all but finitely many points of l.

PROOF. Let $\mathcal{S} \subset l$ be a set of points where P obeys condition Y, and suppose that $|\mathcal{S}| > K(Y) \operatorname{Deg} P$.

Recall from Lemma 13.19 that there is a list of polynomials $Y_j P$ with $\operatorname{Deg} Y_j P \leq C(Y) \operatorname{Deg} P$, for $j = 1, ..., J(Y)$, and that P obeys condition Y at z if and only if the vector $v(Y_j P(z))$ lies in $B_Y \in \{0,1\}^{J(Y)}$.

At each point $z \in \mathcal{S}$, we let $\beta(z) \in B_Y$ be the vector $v(Y_j P(z))$. We let $\mathcal{S}_\beta := \{z \in \mathcal{S} | \beta(z) = \beta\}$. There are at most $2^{J(Y)}$ elements in B_Y, and so by the pigeonhole principle, there is some $\beta \in B_Y$ so that $|\mathcal{S}_\beta| > 2^{-J(Y)} K(Y) \operatorname{Deg} P$. We choose $K(Y) > C(Y) 2^{J(Y)}$, so that

$$|\mathcal{S}_\beta| > C(Y) \operatorname{Deg} P \geq \operatorname{Deg} Y_j P.$$

We fix this value of β. If $\beta_j = 0$, then we see that $Y_j P$ vanishes at $> (\operatorname{Deg} Y_j P)$ points of l. Therefore,$Y_j P$ vanishes on l. If $\beta_j = 1$, then we see that $Y_j P$ fails to vanish at at least one point of l. Therefore, $Y_j P$ vanishes at only finitely many points of l.

Thus at all but finitely many points of l, $v(Y_j P) = \beta \in B_Y$. Hence all but finitely many points of l obey condition Y. □

LEMMA 13.26. Suppose that $Y \subset \mathrm{Poly}_r(\mathbb{C}^3)$ is a constructible condition, for some $r \geq 0$. Then there is a constant $K(Y)$ so that the following holds. Let $P : \mathbb{C}^3 \to \mathbb{C}$ be a polynomial. Suppose that $\mathfrak{L}$ is a set of lines in $\mathbb{C}^3$, and that P obeys Y at all but finitely many points of each line of $\mathfrak{L}$. Suppose that all the lines of $\mathfrak{L}$ are contained in an algebraic surface $Z(Q)$ for an irreducible polynomial Q. If $|\mathfrak{L}| > K(Y) \operatorname{Deg} P \operatorname{Deg} Q$, then P obeys Y at almost every point of $Z(Q)$.

PROOF. For each line $l \in \mathfrak{L}$, we will choose an element $\beta(l) \in B_Y \subset \{0,1\}^{J(Y)}$. Define $\beta_j(l) = 0$ if and only if $Y_j P(z)$ vanishes on l. For all but finitely many points $z \in l$, we have $v(Y_j P(z)) = \beta_j(l)$. We must have $\beta(l) \in B_Y \subset \{0,1\}^{J(Y)}$.

For each $\beta \in B_Y$, we define $\mathfrak{L}_\beta := \{l \in \mathfrak{L} | \beta(l) = \beta\}$. There are at most $2^{J(Y)}$ elements of B_Y, and so by the pigeonhole principle, we can choose $\beta \in B_Y$ so that $|\mathfrak{L}_\beta| \geq 2^{-J(Y)}|\mathfrak{L}|$.

Recall from Lemma 13.19 that $\operatorname{Deg} Y_j P \leq C(Y) \operatorname{Deg} P$. We choose $K(Y) = 2^{J(Y)} C(Y)$, so that

$$|\mathfrak{L}_\beta| > 2^{-J(Y)} K(Y) \operatorname{Deg} P \operatorname{Deg} Q \geq C(Y) \operatorname{Deg} P \operatorname{Deg} Q \geq \operatorname{Deg} Y_j P \operatorname{Deg} Q.$$

Fix this value of β. We consider the behavior of $Y_j P$ on $Z(Q)$ for different values of j. First suppose that $\beta_j = 0$. Then $Y_j P$ vanishes on each line of $\mathfrak{L}_\beta$. So $Z(Y_j P) \cap Z(Q)$ contains all the lines of $\mathfrak{L}_\beta$. Since $|\mathfrak{L}_\beta| > \operatorname{Deg} Y_j P \operatorname{Deg} Q$, the Bezout theorem for lines (Theorem 6.7) implies that Q and $Y_j P$ have a common factor. Since Q is irreducible, it follows that Q divides $Y_j P$, and so $Y_j P$ vanishes on $Z(Q)$. In other words, $v(Y_j P) = \beta_j$ at every point of $Z(Q)$.

On the other hand, suppose that $\beta_j = 1$. Then $Y_j P$ does not vanish on a line $l \in \mathfrak{L}_\beta$ with $l \subset Z(Q)$. In particular we can find at least one point of $Z(Q)$ where $Y_j P$ does not vanish, and so Q does not divide $Y_j P$. In this case, $v(Y_j P) = 1$ for almost every point of $Z(Q)$.

By Lemma 13.23, at almost every point of $Z(Q)$, $v(Y_j P) = \beta_j$ for all j. Hence, at almost every point of $Z(Q)$, P obeys Y. □

In particular, we get the following Corollary.

COROLLARY 13.27. If Y is a constructible condition, then there is a constant $K(Y)$ so that the following holds. Suppose that $P \in \operatorname{Poly}(\mathbb{C}^3)$ is irreducible, and that $Z(P)$ contains greater than $K(Y)(\operatorname{Deg} P)^2$ lines each of which contains greater than $K(Y) \operatorname{Deg} P$ points where P obeys Y. Then almost every point of $Z(P)$ obeys Y.

PROOF. By Lemma 13.25, P obeys Y at almost every point of each of the lines above. Then by Lemma 13.26, P obeys Y at almost every point of $Z(P)$. □

By Proposition 13.18, being (s,r)-flecnodal is a constructible condition. Applying Corollary 13.27 to this condition, we get the following.

COROLLARY 13.28. For any (s,r) there is a constant $K(s,r)$ so that the following holds. Suppose that $P \in \operatorname{Poly}(\mathbb{C}^3)$ is irreducible. Suppose that $\mathfrak{L}$ is a set of lines in $Z(P)$ with $|\mathfrak{L}| > K(s,r)(\operatorname{Deg} P)^2$. Suppose that each line $l \in \mathfrak{L}$ contains greater than $K(s,r) \operatorname{Deg} P$ points that are (s,r)-flecnodal for P. Then almost every point of $Z(P)$ is (s,r)-flecnodal for P.

13.5. From local to global

So far we have studied the local geometry of $Z(P)$. We have studied the (s,r)-flecnodal points for various (s,r). Whether a point $z \in Z(P)$ is (s,r)-flecnodal depends only on the r-jet of P at the point z. We will call such a condition a local condition. In this section we go from local to global. Suppose that every point of a surface $Z(P)$ is r-flecnodal for a large r. Does this imply that $Z(P)$ is actually ruled? The answer is yes, by a classical theorem of Monge, Cayley, and Salmon. See [**Ko**] for some discussion of the history.

THEOREM 13.29. (Cayley-Monge-Salmon) If P is an irreducible polynomial in $\mathrm{Poly}(\mathbb{C}^3)$, and if every point of $Z(P)$ is flecnodal, then $Z(P)$ is ruled.

For our application, we need an analogue of this theorem for doubly ruled surfaces. The doubly ruled case is actually easier.

PROPOSITION 13.30. Suppose that P is an irreducible polynomial in $\mathrm{Poly}(\mathbb{C}^3)$. Suppose that almost every point of $Z(P)$ is doubly flecnodal. Then P has degree 1 or 2, and $Z(P)$ is a plane or regulus.

This Proposition easily implies the classification of doubly ruled surfaces we stated above as Theorem 13.2:

THEOREM. Suppose that $P \in \mathrm{Poly}(\mathbb{C}^3)$ is an irreducible polynomial, and suppose that $Z(P)$ is doubly ruled. Then P has degree 1 or 2, and $Z(P)$ is a plane or regulus.

We will prove Proposition 13.30. We don't give a complete proof of Theorem 13.29, but we introduce many of the ideas. The key issue is the following: given that many points are flecnodal or doubly flecnodal, how do we prove that the surface $Z(P)$ actually contains some lines? We prove this type of result using differential geometry. We will carry out the argument first over $\mathbb{R}$ and then over $\mathbb{C}$. The proof we use here is based on the proof sketched by Kollar in [**Ko**].

Here is a convenient setup for the differential geometry argument. Suppose that $B \subset \mathbb{R}^2$ is a ball, and $h : B \to \mathbb{R}$ is a smooth function. We study the graph of h.

The flecnodal condition involves the second and third derivatives of h. The second derivatives of h form the Hessian, $\nabla^2 h$. Recall that the Hessian of h is a symmetric bilinear form defined as follows: if $v, w \in \mathbb{R}^2$ are vectors, then

$$\nabla_v \nabla_w h := \sum_{i,j} v_i w_j \frac{\partial^2 h}{\partial x_i \partial x_j}. \tag{13.11}$$

There is a similar notation for third derivatives. If $u, v, w \in \mathbb{R}^2$ are vectors, then we define

$$\nabla_u \nabla_v \nabla_w h := \sum_{i,j,k} u_i v_j w_k \frac{\partial^3 h}{\partial x_i \partial x_j \partial x_k}. \tag{13.12}$$

Because partial derivatives commute, these expressions are symmetric: for instance $\nabla_v \nabla_u \nabla_w h = \nabla_w \nabla_v \nabla_u h$. We sometimes abbreviate $\nabla_v^2 h := \nabla_v \nabla_v h$.

A point $(x_1, x_2, h(x_1, x_2))$ is a flecnodal point for the graph of h if and only if it lies in a line which is tangent to the graph of h to third order. This holds if and only if there is a 1-dimensional subspace $K \subset \mathbb{R}^2$ so that for any vector $v \in K$, we have

$$0 = \nabla_v^2 h(x_1, x_2) = \nabla_v^3 h(x_1, x_2). \tag{13.13}$$

Now we are ready to state our local-to-global lemma about graphs $x_3 = h(x_1, x_2)$.

LEMMA 13.31. Suppose that $B \subset \mathbb{R}^2$ is a ball and $h : B \to \mathbb{R}$ is smooth. For each $x \in B$, let $K(x) \subset \mathbb{R}^2$ be a 1-dimensional subspace, depending smoothly on

x. Suppose that for any vector $v \in K(x)$, we have the flecnodal condition

$$\nabla_v^2 h(x) = \nabla_v^3 h(x) = 0.$$

Finally, suppose that $\operatorname{Rank} \nabla^2 h(x) = 2$ at every point $x \in B$.

For each $x_0 \in B$, let $L(x_0) \subset \mathbb{R}^2$ be the line through x_0 with tangent space $K(x_0)$. Then at each point $x \in L(x_0) \cap B$, the space $K(x)$ is equal to the tangent space of $L(x_0)$. As a result, we will show that the restriction of h to each line $L(x_0)$ is linear.

This lemma implies that every point of the graph of h lies in a line segment in the graph of h. (For any $x_0 \in B$, the graph of h over $L(x_0)$ is a line segment in the graph of h containing the point $(x_0, h(x_0))$.)

Before we start the proof, we discuss a subtle point in the statement. Why do we need the hypothesis that $\operatorname{Rank} \nabla^2 h = 2$ at every $x \in B$? If we remove this hypothesis entirely, then the lemma is not true. Suppose that h is a linear function. In this case, $\nabla_v^2 h(x) = \nabla_v^3 h(x) = 0$ for any $x \in B, v \in \mathbb{R}^2$, and $\operatorname{Rank} \nabla^2 h(x) = 0$ for any $x \in B$. Let $K(x) \subset \mathbb{R}^2$ be any 1-dimensional subspace depending smoothly on x. Aside from the hypothesis $\operatorname{Rank} \nabla^2 h = 2$, this example satisfies all the remaining hypotheses of the lemma, no matter what space $K(x)$ we choose. But for a generic choice of K, it will not be true that for every $x \in L(x_0)$, $K(x)$ is the tangent space of x. Lemma 13.31 is also true in the case $\operatorname{Rank} \nabla^2 h(x) = 1$ for all $x \in B$, but this proof is a little more complicated than the proof of Lemma 13.31 – we will discuss it more in the exercises at the end of the chapter.

Proof of Lemma 13.31. Pick a point $x_0 \in B$. We will show that $K(x)$ is constant on $L(x_0)$ in a neighborhood of x_0.

Since $K(x)$ varies smoothly in x, we can find a smooth non-vanishing vector field v on a neighborhood of x_0 so that $v(x) \in K(x)$ for every x. (To do this, pick an affine 1-space A which does not contain zero and is transverse to $K(x_0)$, and then define $v(x)$ to be the intersection point of $K(x)$ with A.)

We let ϕ be the integral curve of v starting at x_0. In other words, ϕ is a map from an interval I to B with $\phi(0) = x_0$ and $\phi'(s) = v(\phi(s)) \in K(\phi(s))$. Since v is a smooth vector field, there is a unique integral curve ϕ by the fundamental theorem about solving ordinary differential equations. The interval I is an open interval around 0.

We claim that $K(\phi(s))$ is constant in s. The proof of this claim is the heart of the lemma. The proof of the claim depends on two facts about $\phi(s)$: $\phi'(s)$ never vanishes and $\phi'(s) \in K(\phi(s))$ for every $s \in I$. Since $0 \neq \phi'(s) \in K(\phi(s))$ and K is 1-dimensional, we see that $K(\phi(s)) = \operatorname{Span}(\phi'(s))$ for every $s \in I$. So to prove that $K(\phi(s)) = \operatorname{Span}(\phi'(s))$ is constant , it suffices to check that $\phi''(s) \in K(\phi(s))$ for every s.

We know that for every s, $\nabla^2_{\phi'(s)} h = \nabla^3_{\phi'(s)} h = 0$. Therefore, we can compute using the Liebniz rule:

$$0 = \frac{d}{ds}\left(\nabla^2_{\phi'(s)} h(\phi(s))\right) = \nabla^3_{\phi'(s)} h(\phi(s)) + 2\nabla_{\phi'(s)} \nabla_{\phi''(s)} h(\phi(s)). \tag{13.14}$$

Since $\nabla^3_{\phi'(s)} h$ vanishes, we conclude that $\nabla_{\phi'(s)} \nabla_{\phi''(s)} h(\phi(s))$ vanishes identically on I. At this moment, we use the fact that $\operatorname{Rank} \nabla^2 h = 2$. For each value of s, we consider the kernel of the linear map

$$w \mapsto \nabla_{\phi'(s)} \nabla_w h(\phi(s)). \tag{13.15}$$

Since $\nabla^2 h$ is non-degenerate, this linear map is onto and its kernel is 1-dimensional. We know that $\nabla^2_{\phi'(s)} h(\phi(s)) = 0$, and so $\phi'(s)$ is in the kernel. We conclude that the kernel is exactly the span of $\phi'(s)$, which is exactly $K(\phi(s))$. Therefore, $\phi''(s) \in K(\phi(s))$ for every s. This finishes the proof of the claim: $K(\phi(s))$ is constant in s.

For all s, $\phi'(s) \in K(\phi(s)) = K(x_0)$. Therefore, $\phi(s)$ stays in the line $L(x_0)$. The image of ϕ must contain a neighborhood of x_0 in the line $L(x_0)$. Therefore, we see that $K(x)$ is constant on a neighborhood of x_0 in $L(x_0)$.

From here, we can quickly prove that $K(x)$ is constant on $L(x_0) \cap B$. Let $A \subset L(x_0) \cap B$ be the set of points x where $K(x) = K(x_0)$. Since $K(x)$ varies continuously in x, the set A is closed. But by the argument in the last paragraph, A is also open. Since B is a ball, $L(x_0) \cap B$ is a convex set, and in particular it is connected. Therefore, $K(x) = K(x_0)$ for every $x \in L(x_0) \cap B$.

We know that $K(x) = K(x_0)$ on $L(x_0) \cap B$. Since $\nabla^2_v h = 0$ for $v \in K$, it follows that the restriction of h to $L(x_0) \cap B$ has zero second derivative. Therefore, h restricted to $L(x_0) \cap B$ is linear. □

We need a complex version of Lemma 13.31. Most of the ideas adapt from $\mathbb{R}$ to $\mathbb{C}$ in a straightforward way, but one or two points require more thought.

Suppose now that $B \subset \mathbb{C}^2$ is a ball, and $h : B \to \mathbb{C}$ is a holomorphic function. We will again study the graph of h.

We first recall how partial derivatives work for a holomorphic function. Suppose that $z_j = x_j + iy_j$ with x_j, y_j real. So x_j, y_j are real coordinates on $\mathbb{C}^2 = \mathbb{R}^4$. The function h is holomorphic if and only if it obeys the Cauchy-Riemann equations: for each j, $\frac{\partial h}{\partial y_j} = i \frac{\partial h}{\partial x_j}$. The Cauchy-Riemann equations assert that the derivative map $dh : \mathbb{C}^2 \to \mathbb{C}$ is complex linear. They can also be written in the following way: for each j,

$$\frac{\partial h}{\partial x_j} + i \frac{\partial h}{\partial y_j} = 0.$$

We note that the graph of h is a complex submanifold if and only if h obeys the Cauchy-Riemann equations. In particular, if $P \in \mathrm{Poly}(\mathbb{C}^3)$ and $O \subset Z(P)$ is an open set of regular points, and if O is the graph of a function $h : B^2 \to \mathbb{C}$, then h will be holomorphic.

Next, recall that the derivative $\frac{\partial h}{\partial z_j}$ is defined as:

$$\frac{\partial h}{\partial z_j} := \frac{\partial h}{\partial x_j} - i \frac{\partial h}{\partial y_j}.$$

Because of the Cauchy-Riemann equations, all of the first derivatives $\frac{\partial h}{\partial x_j}$ and $\frac{\partial h}{\partial j_y}$ can be recovered from $\frac{\partial h}{\partial z_j}$. Higher derivatives of h are defined by iterating. For instance,

$$\frac{\partial^2 h}{\partial z_i \partial z_j} := \frac{\partial}{\partial z_j} \left(\frac{\partial h}{\partial z_i} \right).$$

As usual, partial derivatives commute. The flecnodal condition involves the second and third derivatives of h. The second derivatives of h form the Hessian, $\nabla^2 h$. Recall that the Hessian of h is a symmetric bilinear form defined as follows: if

$v, w \in \mathbb{C}^2$ are vectors, then

$$\nabla_v \nabla_w h := \sum_{i,j} v_i w_j \frac{\partial^2 h}{\partial z_i \partial z_j}. \tag{13.16}$$

There is a similar notation for third derivatives. If $u, v, w \in \mathbb{C}^2$ are vectors, then we define

$$\nabla_u \nabla_v \nabla_w h := \sum_{i,j,k} u_i v_j w_k \frac{\partial^3 h}{\partial z_i \partial z_j \partial z_k}. \tag{13.17}$$

Because partial derivatives commute, these expressions are symmetric: for instance $\nabla_v \nabla_u \nabla_w h = \nabla_w \nabla_v \nabla_u h$. We sometimes abbreviate $\nabla_v^2 h := \nabla_v \nabla_v h$.

A point $(z_1, z_2, h(z_1, z_2))$ is a flecnodal point for the graph of h if and only if it lies in a (complex) line which is tangent to the graph of h to third order. This holds if and only if there is a 1-dimensional (complex) subspace $K \subset \mathbb{C}^2$ so that for any vector $v \in K$, we have

$$0 = \nabla_v^2 h(z_1, z_2) = \nabla_v^3 h(z_1, z_2). \tag{13.18}$$

This point is slightly subtler over $\mathbb{C}$. In particular, suppose that there is a subspace K as above. If we restrict h to the complex line L through (z_1, z_2) tangent to K, then on this line we see that the holomorphic second and third derivatives of h vanish at (z_1, z_2). Now by the Cauchy-Riemann equations, it follows that all the ordinary real second and third derivatives of h on the line L also vanish at (z_1, z_2). Let L' be the complex line in $\mathbb{C}^3$ which is tangent to the graph of h over L at the point $(z_1, z_2, h(z_1, z_2))$. Since the second and third derivatives vanish, L' is tangent to the graph of h to third order.

Now we are ready to state our local-to-global lemma about holomorphic graphs $z_3 = h(z_1, z_2)$.

LEMMA 13.32. Suppose that $h : B \to \mathbb{C}$ is holomorphic, for a ball $B \subset \mathbb{C}^2$. For each $z \in B$, let $K(z) \subset \mathbb{C}^2$ be a 1-dimensional complex subspace, varying smoothly in z. Suppose that for any vector $v \in K(z)$, we have

$$\nabla_v^2 h(z) = \nabla_v^3 h(z) = 0.$$

Finally, suppose that $\operatorname{Rank} \nabla^2 h = 2$ everywhere in B.

For each $z_0 \in B$, let $L(z_0) \subset \mathbb{C}^2$ be the line through z_0 with tangent space $K(z)$. Then at each point $z \in L(z_0) \cap B$, the $K(z)$ is equal to the tangent space of $L(z_0)$. Moreover, the restriction of h to each line $L(z)$ is linear.

PROOF. An important piece of the proof of Lemma 13.31 adapts very readily to the complex case. We consider real curves that are tangent to the subspaces $K(z)$. More precisely, suppose that $\phi : I \to \mathbb{C}^2$ is a curve defined on an interval $I \subset \mathbb{R}$ with $\phi'(s)$ never vanishing, and suppose that $\phi'(s) \in K(\phi(s))$ for every $s \in I$. An important step of the proof is to show that $K(\phi(s))$ is constant in s.

We write $\operatorname{Span} A$ for the complex span of a subset $A \subset \mathbb{C}^2$. Since $K(\phi(s)) = \operatorname{Span} \phi'(s)$, it suffices to check that $\phi''(s) \in K(\phi(s))$ for every $s \in I$.

We know that for every s, $\nabla^2_{\phi'(s)} h = \nabla^3_{\phi'(s)} h = 0$. Therefore, we can compute using the Liebniz rule:

$$0 = \frac{d}{ds} \left(\nabla^2_{\phi'(s)} h(\phi(s)) \right) = \nabla^3_{\phi'(s)} h(\phi(s)) + 2 \nabla_{\phi'(s)} \nabla_{\phi''(s)} h(\phi(s)). \tag{13.19}$$

Since $\nabla^3_{\phi'(s)} h$ vanishes, we conclude that $\nabla_{\phi'(s)} \nabla_{\phi''(s)} h(\phi(s))$ vanishes identically on I. At this moment, we use the fact that $\operatorname{Rank} \nabla^2 h = 2$. At each point s, we consider the kernel of the linear map

$$w \mapsto \nabla_{\phi'(s)} \nabla_w h(\phi(s)). \tag{13.20}$$

Since $\nabla^2 h$ is non-degenerate, this linear map is onto and its kernel is 1-dimensional. We know that $\nabla^2_{\phi'(s)} h(\phi(s)) = 0$, and so $\phi'(s)$ is in the kernel. We conclude that the kernel is exactly the span of $\phi'(s)$, which is exactly $K(\phi(s))$. Therefore, $\phi''(s) \in K(\phi(s))$ for every s.

This finishes the proof of our first claim: if $\phi : I \to \mathbb{C}^2$ is a smooth curve with $\phi'(s)$ never vanishing and $\phi'(s) \in K(\phi(s))$ for all s, then $K(\phi(s))$ is constant in s.

We want to find appropriate curves $\phi(s)$ to apply this claim. Over $\mathbb{R}$, we picked a non-vanishing vector field v with $v(z) \in K(z)$ for every z, and then we let ϕ be an integral curve of z.

We do something similar over $\mathbb{C}$. We fix a point $z_0 \in B$. On a neighborhood of z_0, we choose a non-vanishing smooth vector field v with $v(z) \in K(z)$ for every z. (Again, we can do this by picking an affine 1-dimensional complex space $A \subset \mathbb{C}^2$ transverse to $K(z_0)$ and letting $v(z) = K(z_0) \cap A$.) For each angle θ, we also consider the vector field $e^{i\theta} v$. We let $\phi_\theta(s)$ be the solution to the ordinary differential equation $\phi'_\theta(s) = e^{i\theta} v$ and $\phi_\theta(0) = z_0$. The solution ϕ_θ is defined for s in some interval around 0, until the solution leaves the neighborhood where $v(z)$ is defined. In particular, there is some $\varepsilon > 0$ so that for every θ, the solution is defined on $[0, \varepsilon)$.

By the discussion above, $\phi'_\theta(s) \in K(z_0)$ for every θ, s. Therefore, $\phi_\theta(s) \in L(z_0)$ for every θ, s. Allowing θ to vary in $[0, 2\pi]$ and s to vary in $[0, \varepsilon]$, we can think of θ, s in polar coordinates so that ϕ becomes a map from $B^2(\varepsilon) \subset \mathbb{R}^2$ to $L(z_0)$. Since the solution of an ordinary differential equation depends smoothly on the vector field, ϕ is a smooth map from $B^2(\varepsilon)$ to $L(z_0)$. The map sends 0 to z_0, and its derivative at 0 is an isomorphism. By the inverse function theorem, the image contains an open neighborhood of z_0 in $L(z_0)$. At each point in the image $K(z) = K(z_0)$. Therefore, we see that K is constant on a neighborhood of z_0 in $L(z_0)$.

From here, we can quickly prove that K is constant on $L(z_0) \cap B$. Let $A \subset L(z_0) \cap B$ be the set of points z where $K(z) = K(z_0)$. Since $K(z)$ varies continuously in z, the set A is closed. But by the argument in the last paragraph, A is also open. Since B is a ball, $L(z_0) \cap B$ is a convex set, and in particular it is connected. Therefore, $K(z) = K(z_0)$ for every $z \in L(z_0) \cap B$.

We know that $K(z) = K(z_0)$ on $L(z_0) \cap B$. Since $\nabla^2_v h = 0$ for $v \in K$, it follows that the restriction of h to $L(z_0) \cap B$ has zero (holomorphic) Hessian. Since h is holomorphic, it follows using the Cauchy-Riemann equations that h restricted to $L(z_0) \cap B$ is linear. □

We now have the tools to prove Proposition 13.30. For convenience, we now recall the statement:

PROPOSITION. Suppose that P is an irreducible polynomial in $\operatorname{Poly}(\mathbb{C}^3)$. Suppose that almost every point of $Z(P)$ is doubly flecnodal. Then P has degree 1 or 2, and $Z(P)$ is a plane or regulus.

Remark: Along the way, we will prove that $Z(P)$ contains an open set O where each point is regular and each point lies in two lines in $Z(P)$. So along the way, we will establish Proposition 13.5.

PROOF. By Corollary 13.24, we know that there is an open subset $O \subset Z(P)$ where every point is regular and doubly-flecnodal. By shrinking O, we can assume that it is given by a graph, $z_3 = h(z_1, z_2)$, for a holomorphic function h defined on a small ball $B \subset \mathbb{C}^2$.

If $\nabla^2 h$ vanishes on the whole domain B, then the graph of h is an open subset of a plane. In this case, since P is irreducible, it follows that P is degree 1 and $Z(P)$ is a plane. So we can assume that there is one point of B where $\nabla^2 h$ is not zero. By shrinking the domain B, we can then assume that $\nabla^2 h(z_1, z_2)$ is non-zero on all of B.

Fix $(z_1, z_2) \in B$. Since P is doubly flecnodal at $(z_1, z_2, h(z_1, z_2))$, there must be two different 1-dimensional subspaces $K_1, K_2 \subset \mathbb{C}^2$ so that for any v in K_1 or K_2,

$$0 = \nabla_v^2 h(z_1, z_2) = \nabla_v^3 h(z_1, z_2). \tag{13.21}$$

Now we consider the set of null directions $N := \{v \in \mathbb{C}^2 | \nabla_v \nabla_v h(z_1, z_2) = 0\}$. The expression $\nabla_v \nabla_v h(z_1, z_2)$ is a (non-zero) quadratic polynomial in v. If $\operatorname{Rank} \nabla^2 h(z_1, z_2)$ is 1, then N is a single line. (If the rank of $\nabla^2 h(z_1, z_2) = 1$, then we can make a linear change of coordinates so that $\nabla_v \nabla_v h(z_1, z_2) = v_1^2$. In these coordinates, N is the line $v_1 = 0$.) But we know that there are two different lines where $\nabla_v^2 h = 0$. So for a doubly flecnodal point, $\nabla^2 h$ cannot have rank 1. Therefore, $\operatorname{Rank} \nabla^2 h(z_1, z_2) = 2$ for all $(z_1, z_2) \in B$.

If $\operatorname{Rank} \nabla^2 h(z_1, z_2) = 2$, then the set of null directions N is the union of two lines. (If $\operatorname{Rank} \nabla^2 h(z_1, z_2) = 2$, then we can make a linear change of coordinates so that $\nabla_v \nabla_v h(z_1, z_2) = v_1^2 + v_2^2$. In these coordinates, the null set N is exactly two lines: $v_1 + iv_2 = 0$ and $v_1 - iv_2 = 0$.) Therefore, at each point of B, $K_1(z_1, z_2) \cup K_2(z_1, z_2)$ is exactly $N(z_1, z_2)$.

Since $\nabla^2 h$ is a smooth function on B, and has constant rank 2, it follows that $K_1(z_1, z_2)$ and $K_2(z_1, z_2)$ vary smoothly on B.

We are now in a position to apply Lemma 13.32. For each point $z \in B \subset \mathbb{C}^2$, we let $L_i(z)$ be the line through z tangent to $K_i(z)$. By Lemma 13.32, the function h restricted to each line $L_i(z)$ is linear. We let $\tilde{L}_i(z)$ be the graph of h over $L_i(z)$. We note that $\tilde{L}_i(z)$ is a complex line segment in the graph of h. By abuse of notation, we also let $\tilde{L}_i(z)$ be the complex line in $\mathbb{C}^3$ containing this line segment. We note that $\tilde{L}_i(z)$ is a line in $Z(P)$.

We have now proven that each point in the graph of h lies in two lines in $Z(P)$. The graph of h is an open set $O \subset Z(P)$ where each point is regular and each point lies in two lines in $Z(P)$, establishing Proposition 13.5. Now using all of these lines, we will prove that $Z(P)$ contains arbitrarily many lines in a plane or regulus.

Let $z_0 \in B$. After a linear change of variables in $\mathbb{C}^2$, we can assume that $L_1(z_0)$ is parallel to the z_1-axis and that $L_2(z_0)$ is parallel to the z_2-axis. Now by restricting B to a small ball around z_0, we can assume that $L_j(z)$ is nearly parallel to the z_j-axis for all $z \in B$.

We claim that any point $z \in B$ lies in exactly one line of the form $L_1(w)$. We know that z lies in $L_1(z)$. Suppose that z also lies in $L_1(w)$ for some $w \notin L_1(z)$. But then $\nabla_v^2 h(z)$ must vanish for v in the tangent space of $L_1(z)$ and $L_1(w)$, and

also in the tangent space of $L_2(z)$. But $\nabla_v^2 h(z)$ only vanishes for v in two complex subspaces: $K_1(z) \cup K_2(z)$. By assumption the tangent space of $L_1(w)$ is not $K_1(z)$. The tangent space of $L_1(w)$ is also not $K_2(z)$, because $K_2(z)$ is nearly tangent to the z_2-axis and $L_1(w)$ is nearly tangent to the z_1-axis. This contradiction proves the claim.

Now pick three points in B that are very close to z_0, called w_1, w_2, w_3. Consider the three lines $L_1(w_1), L_1(w_2), L_1(w_3)$. By choosing w_1, w_2, w_3 generically, we can assume that these are three distinct lines. By the last paragraph, the three line segments $L_1(w_m) \cap B$ are disjoint. Recall that $\tilde{L}_1(w_m)$ is the graph of h over $L_1(w_m) \cap B$, which is a line in $Z(P)$. By Proposition 8.14 in the section on reguli, there is a polynomial Q of degree at most 2 so that $Z(Q)$ contains all three lines $\tilde{L}_1(w_m)$.

Now consider many points $u_n \in B$, also very close to z_0. Each line $L_2(u_n)$ must intersect $L_1(w_1), L_1(w_2)$, and $L_1(w_3)$ in B. (Since $L_2(u_n)$ is almost parallel to the z_2-axis, and $L_1(w_m)$ is almost parallel to the z_1-axis, the lines $L_2(u_n)$ and $L_1(w_m)$ must intersect each other. Since u_n and w_m are both close to z_0, the intersection point must also be close to z_0, and so it lies in B.) Since the segments $L_1(w_m) \cap B$ are disjoint, each line $L_2(u_n)$ must intersect $L_1(w_1)$, $L_1(w_2)$, and $L_1(w_3)$ in three distinct points in B. Therefore, the line $\tilde{L}_2(u_n)$ must intersect $\tilde{L}_1(w_1)$, $\tilde{L}_1(w_2)$ and $\tilde{L}_1(w_3)$ in three distinct points. So $\tilde{L}_2(u_n)$ intersects $Z(Q)$ in three distinct points, and so $\tilde{L}_2(u_n)$ lies in $Z(Q)$. Since we can choose infinitely many lines $L_2(u_n)$, there are infinitely many lines in $Z(P) \cap Z(Q)$. By the Bezout theorem for lines, Theorem 6.7, Q and P must have a common factor. Since P is irreducible, P must divide Q, and so the degree of P must be 1 or 2. The surface $Z(Q)$ must be a plane or regulus, and so $Z(P)$ must also be a plane or regulus.

□

13.6. The proof of the main theorem

We now have all the tools we need to prove our main results about $|P_2(\mathfrak{L})|$. We begin with a theorem describing the structure of a configuration of lines where every line has many 2-rich points. This theorem is the heart of the matter.

THEOREM 13.33. There is an absolute constant K so that the following holds. Let $\mathfrak{L}$ be a set of L lines in $\mathbb{C}^3$ so that each line contains $\geq A = KL^{1/2}$ points of $P_2(\mathfrak{L})$. Then $\mathfrak{L}$ lies in $\lesssim L/A$ planes and reguli.

PROOF. Let K be a sufficiently large constant.

Let $\mathfrak{L}$ be a set of L lines in $\mathbb{C}^3$. Suppose that each line of $\mathfrak{L}$ contains $\geq A \geq KL^{1/2}$ points of $P_2(\mathfrak{L})$. We let P be the minimal degree non-zero polynomial that vanishes on $\mathfrak{L}$. By the degree reduction argument (Proposition 11.5), we know that $\operatorname{Deg} P \lesssim L/A$. Choosing K large enough, $\operatorname{Deg} P \leq 10^{-2} L^{1/2}$.

Next we factor P into irreducible factors. We have $P = \prod_j P_j$, where P_j is irreducible. We define $\mathfrak{L}_j \subset \mathfrak{L}$ to be the set of lines that lie in $Z(P_j)$ and don't lie in any other $Z(P_{j'})$. Since P has minimal degree, each $\mathfrak{L}_j$ is non-empty.

LEMMA 13.34. Each line in $\mathfrak{L}_j$ contains $\geq (99/100)A$ points of $P_2(\mathfrak{L}_j)$.

PROOF. Let $l \in \mathfrak{L}_j$. By definition of $\mathfrak{L}_j$, for any $j' \neq j$, $P_{j'}$ does not vanish everywhere on l. So $P_{j'}$ vanishes at $\leq \operatorname{Deg} P_{j'}$ points of l. Therefore, there are $\leq \operatorname{Deg} P$ points of l where $P_{j'}$ vanishes for some $j' \neq j$. But $\operatorname{Deg} P \leq (1/100)A$.

So l contains $\geq (99/100)A$ points of $P_2(\mathfrak{L})$ that don't lie in any other $Z(P_{j'})$. We claim that each of these points lies in $P_2(\mathfrak{L}_j)$. Let x be a point of $l \cap P_2(\mathfrak{L})$, with $P_{j'}(x) \neq 0$ for all $j' \neq j$. The point x lies in at least one other line of $\mathfrak{L}$, l_1. The line l_1 lies in $Z(P)$, but it doesn't lie in $Z(P_{j'})$ for any $j' \neq j$. Therefore, it lies in $Z(P_j)$ and belongs to $\mathfrak{L}_j$. □

Since P is a minimal degree polynomial that vanishes on $\mathfrak{L}$, it follows that P_j is a minimal degree polynomial that vanishes on $\mathfrak{L}_j$. By Proposition 11.5, $\operatorname{Deg} P_j \lesssim |\mathfrak{L}_j|/A \lesssim K^{-1}|\mathfrak{L}_j|^{1/2}$.

To finish the proof, we have to show that each P_j has degree at most 2. Fix a j. For the rest of the proof we study $\mathfrak{L}_j$ and P_j.

Suppose that $z \in P_2(\mathfrak{L}_j)$. We know that z lies in two different lines in $Z(P_j)$. Therefore, z is doubly flecnodal for P_j.

Next we use the contagious vanishing lemmas from Section 13.4. Each line $l \in \mathfrak{L}_j$ contains at least $(99/100)A$ doubly flecnodal points. We know that $(99/100)A \gtrsim K \operatorname{Deg} P_j$. If K is sufficiently large, then by Lemma 13.25, almost every point in each line $l \in \mathfrak{L}_j$ is doubly-flecnodal. The set $\mathfrak{L}_j$ contains $|\mathfrak{L}_j|$ lines. We know that $(\operatorname{Deg} P)^2 \lesssim K^{-2}|\mathfrak{L}_j|$. If K is large enough, then by Lemma 13.26, almost every point of $Z(P_j)$ is doubly flecnodal.

Once we know that almost every point of $Z(P_j)$ is doubly flecnodal we can appeal to the local-to-global arguments. By Proposition 13.30, P_j has degree at most 2.

The number of different factors P_j is $\leq \operatorname{Deg} P \lesssim L/A$. So we conclude that $\mathfrak{L}$ lies in $\lesssim L/A$ algebraic surfaces of degree at most 2.

□

By a standard induction argument, we get the following corollary, which easily implies Theorem 8.3:

COROLLARY 13.35. Suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{C}^3$ that contains at most B lines in any plane or regulus. If $B \geq L^{1/2}$, then

$$|P_2(\mathfrak{L})| \lesssim BL.$$

PROOF. Let K be the constant from Theorem 13.33. Using induction on L, we will prove that

$$|P_2(\mathfrak{L})| \leq KBL.$$

If $|P_2(\mathfrak{L})| \leq KBL$, there is nothing to prove, so we may assume that $|P_2(\mathfrak{L})| > KBL \geq KL^{3/2}$.

We apply Theorem 13.33 with $A = |P_2(\mathfrak{L})|L^{-1} \geq KL^{1/2}$. If each line of $\mathfrak{L}$ contains $\geq A$ points of $P_2(\mathfrak{L})$, then Theorem 11.7 implies that $\mathfrak{L}$ is contained in at most KL/A planes. Therefore, one plane contains at least A/K lines of $\mathfrak{L}$, and so $A/K \leq B$. In this case, we can bound $|P_2(\mathfrak{L})|$ as follows:

$$|P_2(\mathfrak{L})| = AL = K(A/K)L \leq KBL.$$

On the other hand, suppose that there is a line $l \in \mathfrak{L}$ that contains at most A points of $P_3(\mathfrak{L})$. We let $\mathfrak{L}' := \mathfrak{L} \setminus \{l\}$. Now we bound $|P_2(\mathfrak{L})|$ by induction:

$$|P_2(\mathfrak{L})| \leq A + |P_2(\mathfrak{L}')| \leq |P_2(\mathfrak{L})|L^{-1} + KB(L-1).$$

Rearranging we get $\frac{L-1}{L}|P_2(\mathfrak{L})| \leq KB(L-1)$, and so

$$|P_2(\mathfrak{L})| \leq KBL.$$

□

This corollary immediately implies Theorem 13.1:

PROOF. Suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{C}^3$ with at most B lines in any plane or degree 2 surface. We have to show that $|P_2(\mathfrak{L})| \lesssim BL + L^{3/2}$. In particular, $\mathfrak{L}$ contains at most B lines in any plane or regulus. If $B \geq L^{1/2}$, then Corollary 13.35 gives $|P_2(\mathfrak{L})| \lesssim BL$. If $B < L^{1/2}$, then we apply Corollary 13.35 with $B = L^{1/2}$, and we get $|P_2(\mathfrak{L})| \lesssim L^{3/2}$.

□

Finally, we are ready to prove Theorem 8.3.

PROOF. Suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with at most B lines in any plane or degree 2 surface. We have to prove that $|P_2(\mathfrak{L})| \lesssim BL + L^{3/2}$. For each real line $l \subset \mathbb{R}^3$, let $l^{\mathbb{C}}$ be the corresponding complex line in $\mathbb{C}^3$. Let $\mathfrak{L}^{\mathbb{C}}$ be the set of complex lines corresponding to the lines of $\mathfrak{L}$. We note that $|P_2(\mathfrak{L})| \leq |P_2(\mathfrak{L}^{\mathbb{C}})|$.

Next we claim that $\mathfrak{L}^{\mathbb{C}}$ contains at most B lines in any plane or degree 2 surface. Indeed, suppose that P is a complex polynomial of degree 1 or 2 that vanishes on more than B lines of $\mathfrak{L}^{\mathbb{C}}$. Restrict P to $\mathbb{R}^3$, and write $P = P_1 + iP_2$, where P_1, P_2 are real polynomials on $\mathbb{R}^3$ of degree at most 2. The polynomials P_1, P_2 are not both 0, and they both vanish on the lines of $\mathfrak{L}$.

Now we apply Theorem 13.1:

$$|P_2(\mathfrak{L})| \leq |P_2(\mathfrak{L}^{\mathbb{C}})| \lesssim BL + L^{3/2}.$$

□

13.7. Remarks on other fields

In this chapter, we worked over the field of complex numbers. A lot of the material in the chapter can be generalized to other fields.

We should mention first that the main theorem of the chapter is not true over finite fields. A counterexample is given by the Hermitian variety described in Section 3.2. The Hermitian variety H in $\mathbb{F}_q^3$ contains $\sim q^2$ lines, each point of H lies in $\sim q^{1/2}$ lines, and H contains $\sim q^{5/2}$ points. If we randomly select a subset $\mathfrak{L}$ of $100q^{3/2}$ of the lines in H, then $\mathfrak{L}$ will still have $\sim q^{5/2}$ 2-rich points. In other words, $|P_2(\mathfrak{L})| \sim |\mathfrak{L}|^{5/3}$. As we saw in Section 3.2, the Hermitian variety contains $\lesssim q^{1/2}$ lines in any plane. Therefore, our set of lines $\mathfrak{L}$ contains even fewer lines in any plane: $\mathfrak{L}$ contains $\lesssim \log q$ lines in any plane. It is not hard to check that H contains fewer than $q^{1/2}$ lines in any degree 2 surface as well, and so our set of lines $\mathfrak{L}$ contains $\lesssim \log q$ lines in any degree 2 surface. This example shows that Theorem 13.1 does not generalize to $\mathbb{F}_q^3$.

Nevertheless Kollar proved that Theorem 13.1 goes generalize to arbitrary fields $\mathbb{F}$ as long as the number of lines is not too big compared to the characteristic of the field (see Corollary 40 in [**Ko**]). For fields of characteristic p, one requires $|\mathfrak{L}| \leq p^2$. Kollar's theorem applies over all fields, including $\mathbb{C}$. His method has some similar ideas to the method in [**GK2**] and the method here, but also some different ideas.

In connection with other fields, it is worth mentioning that the Chevalley projection theorem, Theorem 13.8 holds over all algebraically closed fields (cf. Theorem 3.16 in [**Ha**]). Also, there is a version of the Cayley-Monge-Salmon theorem, Theorem 13.4, which holds over all fields, provided that the degree of the polynomial is not too large compared to the characteristic – see the discussion before Corollary 40 in [**Ko**].

13.8. Remarks on the bound $L^{3/2}$

In Section 3.6, we raised the question, if $\mathfrak{L}$ is a set of lines in $\mathbb{R}^3$ with at most 10 lines in any plane or degree 2 surface, how many 2-rich points can $\mathfrak{L}$ have? Using ruled surface theory, we proved in this chapter that $|P_2(\mathfrak{L})| \lesssim L^{3/2}$. In Chapter 10, we worked on the same problem using polynomial partitioning. Under a somewhat stronger hypothesis, Theorem 10.1 gives the bound $|P_2(\mathfrak{L})| \leq C(\varepsilon)L^{(3/2)+\varepsilon}$ for any $\varepsilon > 0$. These two methods are very different from each other, but they lead to essentially the same bound. On the other hand, I don't know of any example where $|P_2(\mathfrak{L})|$ is close to $L^{3/2}$. In this section, we briefly discuss why each of these two methods does not do better than $L^{3/2}$.

The method in this chapter is based on studying the lowest degree polynomial that vanishes on the set of lines $\mathfrak{L}$. Suppose that $\mathfrak{L}$ has the following structure: there are $L^{.51}$ planes or reguli S_j and each surface S_j contains $L^{.49}$ lines of $\mathfrak{L}$. The number of 2-rich points of $\mathfrak{L}$ is $\sim L^{.51}(L^{.49})^2 = L^{1.49}$. Suppose we consider the lowest degree polynomial P that vanishes on $\mathfrak{L}$. Multiplying together the polynomials defining the surfaces S_j gives a polynomial P_{prod} of degree $\sim L^{.51}$. But this polynomial is not the lowest degree polynomial vanishing on $\mathfrak{L}$. By parameter counting, we know that there is a non-zdero polynomial P vanishing on $\mathfrak{L}$ with degree $\lesssim L^{1/2}$. The zero set of P probably does not contain any planes or reguli, and so it is not clear how this polynomial P could be helpful in the problem.

A second observation is that if we consider lines in finite fields, then the bound $L^{3/2}$ is essentially sharp. Suppose that $\mathfrak{L}$ is a set of q^2 randomly chosen lines in $\mathbb{F}_q^3$. There are a total of $\sim q^4$ lines in $\mathbb{F}_q^3$, and each point lies in $\sim q^2$ of these lines. Therefore, the probability that a given point lies in at least two lines of $\mathfrak{L}$ is ~ 1. So with high probability, $|P_2(\mathfrak{L})| \sim q^3$. On the other hand, every plane in $\mathbb{F}_q^3$ contains $\sim q^2$ lines. So for a fixed plane $\pi \subset \mathbb{F}_q^3$, the probability that π contains at least B lines of $\mathfrak{L}$ is $\lesssim e^{-cB}$. The total number of planes in $\mathbb{F}_q^3$ is $\sim q^3$. Therefore, with high probability, every plane in $\mathbb{F}_q^3$ contains $\lesssim \log q$ lines of $\mathfrak{L}$. Every irreducible degree 2 surface in $\mathbb{F}_q^3$ contains $\lesssim q$ lines of $\mathfrak{L}$. Therefore, the probability that a given degree 2 surface contains more than 10 lines of $\mathfrak{L}$ is $\lesssim q^{-10}$. On the other hand, the number of irreducible degree 2 surfaces in $\mathbb{F}_q^3$ is $\lesssim q^9$, because $\operatorname{Dim}\operatorname{Poly}_2(\mathbb{F}_q^3) = 10$. Therefore, with high probability, every regulus in $\mathbb{F}_q^3$ contains at most 10 lines of $\mathfrak{L}$.

Exercise 13.1. Consider βq^2 random lines $\mathfrak{L}$ in $\mathbb{F}_q^3$ for a parameter β. Show that for any $\varepsilon > 0$, there exists a constant $B(\varepsilon)$, so that for all q, there exists a set of lines $\mathfrak{L}$ in $\mathbb{F}_q^3$ with $|P_2(\mathfrak{L})| \gtrsim L^{(3/2)-\varepsilon}$ and so that every plane or regulus in $\mathbb{F}_q^3$ contains at most $B(\varepsilon)$ lines of $\mathfrak{L}$.

This example in finite fields does not easily generalize to lines in $\mathbb{R}^3$. Based on our experience with the Szemerédi-Trotter theorem, it is reasonable to try to use a partitioning argument to do better. We tried this approach in Chapter 10, but we were not able to do any better. In fact, we got a slightly weaker bound. Let

us explain why the polynomial partitioning method also got stuck at the exponent $3/2$. Suppose that $\mathfrak{L}$ is a set of L lines in $\mathbb{R}^3$ with KL^α 2-rich points for some parameters K, α. Suppose that we do a degree D polynomial partitioning, where D is a parameter that we can choose, and suppose that all the 2-rich points end up in the cells. We have $\sim D^3$ cells O_i. Each cell contains $\sim D^{-3}KL^\alpha$ 2-rich points of $\mathfrak{L}$. Let $\mathfrak{L}_i \subset \mathfrak{L}$ be the set of lines that intersect the cell O_i. We know that $\sum_i |\mathfrak{L}_i| \lesssim DL$, and so for most cells O_i, $|\mathfrak{L}_i| \sim D^{-2}L$. We would like to understand whether this divide-and-conquer approach is making progress. To do that, we write the number of 2-rich points of $\mathfrak{L}_i$ in terms of $|\mathfrak{L}_i|$:

$$|P_2(\mathfrak{L}_i)| \gtrsim D^{-3}KL^\alpha \sim D^{-3}K(D^2|\mathfrak{L}_i|)^\alpha = D^{2\alpha-3}K|\mathfrak{L}_i|^\alpha.$$

For comparison, we recall that

$$|P_2(\mathfrak{L})| = K|\mathfrak{L}|^\alpha.$$

These equations have the same form, and the constant K on the whole space is replaced by $D^{2\alpha-3}K$ for the lines in a typical cell. If $\alpha > 3/2$, then $2\alpha - 3 > 0$, and we see that the intersection pattern in each cell is more extreme than the original pattern. This provides a good basis for doing induction. On the other hand, if $\alpha < 3/2$, then we see that the intersection pattern in each cell is less extreme than the original pattern. In this situation, it is not clear how such a cell decomposition could be helpful.

13.9. Exercises related to projection theory

In this section, we give some exercises related to projection theory and flecnodes.

EXERCISE 13.2. Let $\mathrm{Flec}(k-\mathrm{plane})_r \subset \mathrm{Poly}_r(\mathbb{C}^n)$ be the set of polynomials that vanish on some k-plane through 0. Prove that $\mathrm{Flec}(k-\mathrm{plane})_r$ is a constructible set.

EXERCISE 13.3. Note that $\mathrm{Flec}(2-\mathrm{plane})_2 \subset \mathrm{Poly}_r(\mathbb{C}^3)$ is closely related to flat points. If $P \in \mathrm{Poly}(\mathbb{C}^3)$, a point $z \in Z(P) \subset \mathbb{C}^3$ is a flat point if and only if z is a regular point and $J^2P(z) \in \mathrm{Flec}(2-\mathrm{plane})_2$.

This point of view can be used to give an alternate treatment of the study of critical and flat points from Section 11.6. This treatment works over $\mathbb{C}$, and it can be used to prove that Theorem 11.1 holds for complex lines in $\mathbb{C}^3$.

EXERCISE 13.4. If $\mathfrak{L}$ is a set of N^2 complex lines in $\mathbb{C}^3$ with at most N lines in any complex 2-plane in $\mathbb{C}^3$, and if X is a finite set with at least N points of X on every line of $\mathfrak{L}$, then prove that $|X| \gtrsim N^3$.

This result contrasts with an example using thin tubular neighborhoods of complex lines in Section 15.9.

EXERCISE 13.5. Let $Y \subset \mathbb{C}^n$ be a constructible set. Recall that we say that something occurs at almost every point of $\mathbb{C}^n$ if the set of exceptional points is contained in $Z(P)$ for a non-zero polynomial $P \in \mathrm{Poly}(\mathbb{C}^n)$. Prove that either Y contains almost every point of $\mathbb{C}^n$ or else the complement of Y contains almost every point of $\mathbb{C}^n$.

In the next set of exercises, we use Chevalley's projection theorem, Theorem 13.8, to explore some basic facts about constructible sets in $\mathbb{C}^n$. In particular, we

define the dimension and the degree of constructible sets in $\mathbb{C}^n$ and show their basic properties. These exercises have to do with how a constructible set intersects k-planes in $\mathbb{C}^n$.

If $b \in \mathbb{C}^{n-k}$, and m is a linear map from $\mathbb{C}^k$ to $\mathbb{C}^{n-k}$, then we let $\Pi(m, b)$ denote the k-plane

$$\Pi(m, b) := \{(x, mx + b) \in \mathbb{C}^n | x \in \mathbb{C}^k\}.$$

We let $G_{k,n}$ be the set of all possible (m, b), where $b \in \mathbb{C}^{n-k}$, and m is a linear map from $\mathbb{C}^k$ to $\mathbb{C}^{n-k}$. Note that we can identify $G_{k,n}$ with $\mathbb{C}^{(k+1)(n-k)}$. Almost every k-plane in $\mathbb{C}^n$ can be written in the form $\Pi(m, b)$, although not every k-plane. The set G_k is a convenient approximation of the affine Grassmannian of $\mathbb{C}^n$ - the set of all k-planes in $\mathbb{C}^n$.

EXERCISE 13.6. Let $Y \subset \mathbb{C}^n$ be a constructible set. Prove that the set of k-planes which intersect Y is constructible:

$$\{(m, b) \in G_{k,n} \text{ so that } \Pi(m, b) \cap Y \text{ is non-empty}\} \text{ is constructible.}$$

Combining Exercise 13.6 and Exercise 13.5, we see that either almost every k-plane intersects Y, or almost every k-plane does not intersect Y. The codimension of Y, $\mathrm{CoDim}(Y)$, is defined to be the smallest k so that almost every k-plane intersects Y. The dimension of Y is defined to be $\mathrm{Dim}(Y) = n - \mathrm{CoDim}(Y)$.

EXERCISE 13.7. Suppose that $L : \mathbb{C}^n \to \mathbb{C}^n$ is an invertible linear map. Suppose that $Y \subset \mathbb{C}^n$ and $Y' = L(Y)$. Prove that $\mathrm{Dim}(Y) = \mathrm{Dim}(Y')$.

EXERCISE 13.8. If $Y \subset \mathbb{C}^n$ is an infinite constructible set, show that almost every $n - 1$-plane intersects Y. Conclude that $\mathrm{Dim}(Y) \geq 1$.

EXERCISE 13.9. Let $Y \subset \mathbb{C}^n$ be a constructible set with $\mathrm{CoDim}(Y) = k$. By definition, almost every $(k-1)$-plane $\Pi \in G_{k-1,n}$ is disjoint from Y.

Prove that for almost every k-plane $\Pi' \in G_{k,n}$, almost every $(k-1)$-plane $\Pi \subset \Pi'$ is disjoint from Y.

Combining this observation with Exercise 13.8, prove that for almost every k-plane $\Pi' \in G_{k,n}$, $Y \cap \Pi'$ is finite (and non-zero).

Conclude that if $k = \mathrm{CoDim}(Y)$, then almost every k-plane intersects Y in finitely many points.

EXERCISE 13.10. Let $Y \subset \mathbb{C}^n$ be a constructible set. For any integer s, prove that the set of k-planes which intersect Y in exactly s points is constructible:

$$\{(m, b) \in G_{k,n} \text{ so that } |\Pi(m, b) \cap Y| = s\} \text{ is constructible.}$$

EXERCISE 13.11. A countable union of constructible sets is not necessarily constructible. Nevertheless, if $Y_j \subset \mathbb{C}^N$ are constructible sets, and the countable union $\cup_{j=1}^{\infty} Y_j$ contains almost every point of $\mathbb{C}^N$, prove that one of the Y_j contains almost every point of $\mathbb{C}^N$.

EXERCISE 13.12. Let $Y \subset \mathbb{C}^n$ be a constructible set. Let $k = \mathrm{CoDim}(Y)$. Prove that there exists an integer $D \geq 1$ so that almost every k-plane intersects Y in a set of cardinality D. The integer D is called the degree of Y.

13.10. Exercises related to differential geometry

The proof of the incidence theorem in this chapter used algebraic geometry in a crucial way, and it also used some differential geometry. In this section, we do some exercises related to the differential geometry tools in the chapter. The goal of the exercises is to understand some geometric facts related to the shapes you can make with a piece of paper. If you have time, try taking a piece of paper and see what shapes you can make with it, without folding or crumpling. In some ways, it seems pretty flexible. For instance, it is easy to roll the paper into a thin cylinder or cone. On the other hand, without folding or crumpling, it is hard to get the paper to fit into a small ball. You can roll it up into a narrow tube, but the shape always seems to be long in one direction. If you look carefully at it, you may see that the surface contains a lot of line segments. The presence of these line segments forces the shape to have at least one long direction. This phenomenon was established by Darboux and others in the late 19th century. See [**CL**] or [**JP**] for more recent references. Over the exercises, we will explore why this happens and say it in a more precise way. The tools are similar to the ones we used in Section 13.5.

Suppose that $\Sigma \subset \mathbb{R}^3$ is a smooth 2-dimensional surface. (We use the word smooth to mean C^∞.) Near a point $p \in \Sigma$, we can choose orthonormal coordinates x_1, x_2, x_3 so that Σ can be written locally as a graph:

$$x_3 = h(x_1, x_2).$$

In the arguments in Section 13.5, the rank of the Hessian, $\operatorname{Rank} \nabla^2 h$, played an important role. We begin by exploring the geometric meaning of this rank. Our first observation is that the rank of the Hessian at a point $p \in \Sigma$ does not depend on the choice of orthonormal coordinates (x_1, x_2, x_3). The rank only depends on the geometry of the surface Σ. We state this as an exercise.

Exercise 13.13. Suppose that $R : \mathbb{R}^3 \to \mathbb{R}^3$ is a rotation, or more generally a rigid motion. Suppose that $\Sigma \subset \mathbb{R}^3$ is a smooth 2-dimensional submanifold, let $\tilde{\Sigma} = R(\Sigma)$ and let $\tilde{p} = R(p)$. Suppose that near p, Σ is described by a graph $x_3 = h(x_1, x_2)$, and near $\tilde{p}$, $\tilde{\Sigma}$ is described as a graph $\tilde{x}_3 = \tilde{h}(\tilde{x}_1, \tilde{x}_2)$. Show that $\operatorname{Rank} \nabla^2 h(p_1, p_2) = \operatorname{Rank} \nabla^2 \tilde{h}(\tilde{p}_1, \tilde{p}_2)$.

A similar statement holds for m-dimensional submanifolds of $\mathbb{R}^n$ for any m, n.

To study the geometry of a surface Σ near a point $p \in \Sigma$, it is convenient to choose orthonormal coordinates so that $p = 0$ and the tangent plane of Σ at p is the (x_1, x_2)-plane. In terms of the function h, this means that $0 = h(0) = \nabla h(0)$. (This choice of coordinates is helpful in the exercise above, for example.) In these coordinates, h has the form

$$h(x_1, x_2) = Q(x_1, x_2) + O(|x|^3),$$

where Q is a homogeneous polynomial of degree 2, which we can think of as a symmetric matrix A. The precise relationship between Q and A is that

$$Q(x) = x^t A x.$$

Here $x = (x_1, x_2) \in \mathbb{R}^2$ is a vector and x^t denotes the transpose of x. The eigenvalues of the matrix A are called the principal curvatures of Σ at p, and the matrix A itself is the second fundamental form of Σ at p.

We see from this discussion that $\operatorname{Rank} \nabla^2 h \leq 1$ at a point of Σ if and only if one of the principal curvatures of Σ vanishes at p. Now the Gauss curvature of Σ at

p is the product of the two principal curvatures, and so we see that $\operatorname{Rank} \nabla^2 h \leq 1$ at a point $p \in \Sigma$ if and only if the Gauss curvature of Σ vanishes at p.

The Gauss curvature is one of the main characters in the differential geometry of surfaces in $\mathbb{R}^3$. One of the important properties of the Gauss curvature is that it is invariant under isometries. Suppose that $\Sigma_0 \subset \mathbb{R}^3$ is a smooth 2-dimensional surface. Recall that a smooth embedding $\phi : \Sigma_0 \to \mathbb{R}^3$ is called an isometric embedding if ϕ preserves the lengths of tangent vectors: for any $p \in \Sigma_0$ and $v \in T_p\Sigma_0$, $|d\phi(v)| = |v|$. If Σ is the image of the isometric embedding $\phi : \Sigma_0 \to \mathbb{R}^3$, then the Gauss curvature of Σ_0 at p is the same as the Gauss curvature of Σ at $\phi(p)$. Gauss called this result the Theorem Egregium. For an explanation of this result, see [**Ca**].

If Σ_0 is an open subset of a plane, then the principal curvatures of Σ_0 vanish at every point, and so Σ_0 has zero Gaussian curvature. If $\phi : \Sigma_0 \to \mathbb{R}^3$ is an isometric embedding with image Σ, then the Theorem Egregium implies that Σ has zero Gaussian curvature. Near any point, we can write Σ as a graph, $x_3 = h(x_1, x_2)$, and by the discussion above we see that $\operatorname{Rank} \nabla^2 h \leq 1$ at all points.

The fact that $\operatorname{Rank} \nabla^2 h \leq 1$ has some important consequences. We start by discussing the algebraic consequences and then we work towards the geometric consequences.

Exercise 13.14. Suppose that $\operatorname{Rank} \nabla^2 h(x) \leq 1$. If $\nabla_v^2 h(x) = 0$, then prove that $\nabla_w \nabla_v h(x) = 0$ for any vector w.

If $\operatorname{Rank} \nabla^2 h(x) = 1$, and if v is a non-zero vector with $\nabla_v^2 h(x) = 0$, then prove that

$$\nabla_{w_1} \nabla_{w_2} h(x) = 0 \text{ if and only if } w_1 \text{ or } w_2 \text{ is in } \operatorname{Span}(v). \tag{13.22}$$

Suppose that $h : B \to \mathbb{R}$ is a smooth function defined on a ball $B \subset \mathbb{R}^2$ obeying the condition $\operatorname{Rank} \nabla^2 h \leq 1$. We are going to explore the geometric consequences of this condition. Suppose that v is a smooth vector-field on B which obeys the condition

$$\nabla_v^2 h(x) = 0 \text{ for all } x \in B.$$

Then surprisingly, v is forced to obey the stronger condition

$$\nabla_v^3 h(x) = 0 \text{ for all } x \in B. \tag{13.23}$$

Exercise 13.15. Prove Equation 13.23. To start, observe that since $\nabla_v^2 h(x) = 0$ on all of B, we can differentiate to get

$$0 = \partial_v (\nabla_v^2 h(x)) \text{ for all } x \in B.$$

Expand out the left-hand side and use Exercise 13.14 to find $\nabla_v^3 h(x)$.

For the rest of this section, we focus on the special case that $\operatorname{Rank} \nabla^2 h(x) = 1$ for all $x \in B$. In this special case, the geometry is cleaner and simpler. In this case, for any $x \in B$, there is a 1-dimensional subspace $K(x)$ of vectors v so that $\nabla_v^2 h(x) = 0$. This 1-dimensional subspace $K(x)$ depends smoothly on x. In the neighborhood of a point $x_0 \in B$, we can always find a smooth non-vanishing vector field v with $v(x) \in K(x)$ for all x. (See the beginning of the proof of Lemma 13.31 for a construction of v.) By Equation 13.23, we see that every point in the graph of h is flecnodal! This is really a little surprising because the condition $\operatorname{Rank} \nabla^2 h = 1$ is a condition about the second derivatives of h, but being flecnodal is about third derivatives. Nevertheless it is true.

Now a variation of the Cayley-Monge-Salmon theorem implies that every point in the graph of h lies in a line segment in the graph of h. The variation that we need here is the version of Lemma 13.31 when $\text{Rank}\,\nabla^2 h = 1$ instead of $\text{Rank}\,\nabla^2 h = 2$. Here is the statement.

LEMMA 13.36. Suppose that $h : B \to \mathbb{R}$ is a smooth function, for a ball $B \subset \mathbb{R}^2$. For each $x \in B$, let $K(x) \subset \mathbb{R}^2$ be a 1-dimensional subspace, varying smoothly in x. Suppose that for any vector $v \in K(x)$, we have

$$\nabla^2_v h(x) = \nabla^3_v h(x) = 0.$$

Finally, suppose that $\text{Rank}\,\nabla^2 h = 1$ everywhere in B.

For each $x_0 \in B$, let $L(x_0)$ be the line through x_0 with tangent space $K(x)$. Then at each point $x \in L(x_0) \cap B$, the $K(x)$ is equal to the tangent space of $L(x_0)$. Moreover, the restriction of h to each line $L(x)$ is linear.

There is also a complex version of this lemma, where we replace $\mathbb{R}$ with $\mathbb{C}$ and assume that h is holomorphic. The proof is essentially the same in the real or complex cases.

The proof of the Cayley-Monge-Salmon theorem about ruled surfaces requires both the rank 1 case and the rank 2 case. See [**Ko**] for an outline of the proof, explaining how to boil the proof down to these two cases. Lemma 13.36 is probably the trickiest step in the proof of Cayley-Monge-Salmon. We will explain the main steps of the proof in the exercises below.

Lemma 13.36 implies a lot about the geometry of isometric embeddings of a planar surface, at least in the special case that the rank of the Hessian is identically 1.

EXERCISE 13.16. Suppose that ϕ is a smooth isometric embedding from the unit disk into $\mathbb{R}^3$ with image Σ. Suppose in addition that at each point of Σ', the second fundamental form has rank exactly 1. (Therefore, if Σ is locally given by a graph $x_3 = h(x_1, x_2)$, then $\text{Rank}\,\nabla^2 h$ is identically 1.)

Using Lemma 13.36, prove that every point of Σ lies in a line segment in Σ, and the endpoints of this line segment lie in $\partial\Sigma$. Using this, prove that Σ contains a line segment of length 2, and conclude that Σ is not contained in any ball of radius < 1.

The proof of Lemma 13.36 involves similar ideas to the proof of Lemma 13.31 (the rank 2 case), but the proof is trickier and more complicated. Before diving into the proof, let us talk a little about what makes the proof trickier. Suppose that $h : B \to \mathbb{R}$ is a smooth function and ϕ is a smooth map from an interval I into B. We say that ϕ is a flecnodal curve if

- For each $s \in I$, $\phi'(s) \neq 0$.
- For each $s \in I$, $\nabla^2_{\phi'(s)} h(\phi(s)) = \nabla^3_{\phi'(s)} h(\phi(s)) = 0$.

In the proof of Lemma 13.31, we showed that if $\text{Rank}\,\nabla^2 h(x) = 2$ for all $x \in B$, then every flecnodal curve is a straight line. This step was the hardest part of the proof of the Lemma. Using the hypothesis that at every point there is a flecnocal direction, it was straightforward to construct a flecnodal curve through every point of the domain B. Then we showed that each of these flecnodal curves was a straight line. However, for a general smooth h it is not true that every flecnodal curve is a straight line. We give a counterexample in the next exercise.

EXERCISE 13.17. Suppose that h is the smooth function

$$h(x_1, x_2) = (x_1^2 + x_2^2 - 1)^2.$$

Let $\phi : \mathbb{R} \to \mathbb{R}^2$ be given by $\phi(s) = (\cos s, \sin s)$. Check that ϕ is a flecnodal curve. The image of ϕ is the unit circle and not a straight line.

In this example, the rank of $\nabla^2 h(x)$ is sometimes 1 and sometimes 2. In particular, the rank of $\nabla^2 h(x) = 1$ for x in the unit circle. This example is not, however, a counterexample to Lemma 13.36. The circle is a flecnodal curve, and at every point in the circle, there is a flecnodal direction. But for other values of x, there is no flecnodal direction. Under the assumptions of Lemma 13.36, it is not hard to show that there is a flecnodal curve through every point of the domain. Using this whole family of curves, instead of just one curve, we will show that all of the curves are straight lines. The need to use a whole family of flecnodal curves instead of a single flecnodal curve is the new wrinkle in the proof of Lemma 13.36.

We give an outline of the proof, leaving the calculations as exercises for the reader.

OUTLINE OF THE PROOF OF LEMMA 13.36. Let $\Phi(r, s)$ be a map from $(-\varepsilon, \varepsilon)^2$ to $B \subset \mathbb{C}^2$ so that at each point (r, s),

$$0 \neq \frac{\partial \Phi}{\partial s}(r, s) \in K(\Phi(r, s)).$$

To find such a function, we may first pick a non-vanishing vector field v on B with $v(x) \in K(x)$. Then we define Φ on $(-\epsilon, \epsilon) \times \{0\}$. For each $r \in (-\epsilon, \varepsilon)$, we solve the ordinary differential equation $\frac{\partial \Phi}{\partial s}(r, s) = v(\Phi(r, s))$. In this way, we define $\Phi(r, s)$ on the whole domain $(-\varepsilon, \varepsilon)^2$. We can think of $\Phi(r, s)$ as a 1-parameter family of integral curves of v – hence as a 1-parameter family of flecnodal curves.

Since $\frac{\partial \Phi}{\partial s}(r, s) \in K(\Phi(r, s))$, we know that

$$\nabla^2_{\frac{\partial \Phi}{\partial s}} h(\Phi) = 0. \tag{13.24}$$

and

$$\nabla^3_{\frac{\partial \Phi}{\partial s}} h(\Phi) = 0. \tag{13.25}$$

By choosing the initial curve $\Phi(r, 0)$ in a generic way, we can assume that $\frac{\partial \Phi}{\partial r}(0, 0) \notin K(\Phi(0, 0))$. After possibly shrinking the domain of Φ, we can then assume that for all (r, s),

$$\frac{\partial \Phi}{\partial r}(r, s) \notin K(\Phi(r, s)). \tag{13.26}$$

We would like to prove that $\frac{\partial^2 \Phi}{\partial s^2}(r, s) \in K(\Phi(r, s))$ for every (r, s), which implies that $K(\Phi(r, s))$ is constant in the s variable. This implies that each curve $s \mapsto \Phi(r, s)$ lies in the line $L_{\Phi(r,0)}$, showing that each flecnodal curve in our family is a straight line.

We will repeatedly use the fact that $\operatorname{Rank} \nabla^2 h = 1$ via Lemma 13.14. Lemma 13.14 tells us that

$$\nabla_{w_1} \nabla_{w_2} h(x) = 0 \text{ if and only if at least one of } w_1, w_2 \text{ lies in } K(x). \tag{13.27}$$

We will eventually calculate that $\nabla_{\frac{\partial\Phi}{\partial r}}\nabla_{\frac{\partial^2\Phi}{\partial s^2}}h = 0$. Since $\frac{\partial\Phi}{\partial r}$ is not in K, Equation 13.27 implies that $\frac{\partial^2\Phi}{\partial s^2} \in K$ as desired. We build up to calculating $\nabla_{\frac{\partial\Phi}{\partial r}}\nabla_{\frac{\partial^2\Phi}{\partial s^2}}h$ in a sequence of steps, combining the equations we have found so far.

Since $\frac{\partial\Phi}{\partial s} \in K$, Equation 13.27 tells us that for any vector w:

$$\nabla_w \nabla_{\frac{\partial\Phi}{\partial s}} h(\Phi) = 0. \tag{13.28}$$

In particular,

$$\nabla_{\frac{\partial\Phi}{\partial r}} \nabla_{\frac{\partial\Phi}{\partial s}} h(\Phi) = 0. \tag{13.29}$$

Differentiating Equation 13.24 with respect to r, we get

$$0 = \frac{\partial}{\partial r}\left(\nabla^2_{\frac{\partial\Phi}{\partial s}} h(\Phi)\right).$$

Exercise 13.18. Expand the right-hand side and use Equation 13.28 to show that

$$\nabla_{\frac{\partial\Phi}{\partial r}} \nabla^2_{\frac{\partial\Phi}{\partial s}} h(\Phi) = 0. \tag{13.30}$$

Next we differentiate Equation 13.29 with respect to s, to get

$$0 = \frac{\partial}{\partial s}\left(\nabla_{\frac{\partial\Phi}{\partial r}}\nabla_{\frac{\partial\Phi}{\partial s}} h(\Phi)\right).$$

Exercise 13.19. Expand the right-hand side and use Equations 13.28 and 13.30 to show that

$$\nabla_{\frac{\partial\Phi}{\partial r}} \nabla_{\frac{\partial^2\Phi}{\partial s^2}} h(\Phi) = 0. \tag{13.31}$$

By Equation 13.27, this equation implies that

$$\frac{\partial^2\Phi}{\partial s^2} \in K. \tag{13.32}$$

The rest of the proof of Lemma 13.36 is the same as for Lemma 13.31.

□

In this discussion, we have developed a pretty good understanding of isometric embeddings of the unit disk into $\mathbb{R}^3$, in the special case that the second fundamental form of the image has rank identically 1. For the reader interested in differential geometry, it could be a good project to try to understand which features of our discussion extend to general smooth isometric embeddings.

https://doi.org/10.1090//ulect/064/14

CHAPTER 14

The polynomial method in differential geometry

In this chapter, we discuss some ideas from differential geometry that are analogous to the ideas we have seen in combinatorics and in coding theory. We consider the question, "what is special about polynomials?" from the point of view of differential geometry. Eventually, using these observations, we will prove a differential geometry theorem with no apparent connection to polynomials. The proof of this result is analogous to the proof of the finite field Nikodym theorem - based on parameter counting and on the vanishing lemma.

From the point of view of differential geometry, polynomials are strikingly efficient. There are a lot of interesting examples of this efficiency in Arnold's essay "Topological economy principle in algebraic geometry", [**Ar**]. We begin with a short survey of results about complex polynomials. These results are interesting, but the proofs are not closely related to the methods in this book.

Next we switch from complex to real polynomials. In particular, we discuss a recent theorem of Gromov on the efficiency of the space $\mathrm{Poly}_D(\mathbb{R}^n)$. The proof uses the Stone-Tukey ham sandwich theorem, which plays an important role in the polynomial partitioning arguments in Chapter 10. This argument from differential geometry helped to suggest the polynomial partitioning arguments there.

This chapter requires a little background in differential geometry and topology. A good reference is the book *Differential Topology* [**GP**]. The chapter is not used anywhere else in the book, so it would be fine to skip it, or to read it for the main ideas without following every detail.

14.1. The efficiency of complex polynomials

We recall the definition of regular points and regular values from differential topology. Suppose that $f : \mathbb{R}^m \to \mathbb{R}^n$ is a smooth map. A point $x \in \mathbb{R}^m$ is called a regular point if the derivative $df_x : \mathbb{R}^m \to \mathbb{R}^n$ is surjective. A point $y \in \mathbb{R}^n$ is called a regular value if every point $x \in f^{-1}(y)$ is a regular point.

If $f : \mathbb{R}^n \to \mathbb{R}^n$ is a smooth map between spaces of the same dimension, and if x is a regular point, then df_x is either orientation preserving or orientation reversing. We define $\mu_f(x) = +1$ if $\det df_x > 0$ and $\mu_f(x) = -1$ if $\det df_x < 0$.

Another fundamental object from topology is the winding number. If $f : S^1 \to \mathbb{R}^2$, then we write $W(f, 0)$ for the winding number of f around 0. (For an introduction to winding numbers see Chapter 3 of [**Ful**] or page 86 of [**GP**].)

Now there is a fundamental theorem of differential topology that connects the winding number of a map f and the multiplicities $\mu_f(x)$ for points $x \in Z(f)$.

THEOREM 14.1. (Page 87 of [**GP**]) If $F : \bar{B}^2 \to \mathbb{R}^2$ is a smooth map and 0 is a regular value, and if $F|_{S^1}$ does not vanish, then

$$W(F|_{S^1}, 0) = \sum_{x \in Z(F) \cap B^2} \mu(x).$$

We will use this theorem to study a complex polynomial in one variable. We will prove that, in a certain sense, a complex polynomial has as few zeroes as possible. We identify $\mathbb{C}$ with $\mathbb{R}^2$.

THEOREM 14.2. Suppose that $P : \mathbb{C} \to \mathbb{C}$ is a complex polynomial, and suppose that $F : \mathbb{R}^2 \to \mathbb{R}^2$ is a smooth function that agrees with P outside of the unit disk. Also suppose that 0 is a regular value for both F and P. Then $|Z(P)| \leq |Z(F)|$.

PROOF. We can assume that P does not vanish on the unit circle. (If P vanishes on the unit circle, just replace the unit disk by a disk of radius $(1 + \varepsilon)$ and rescale the domain.) By hypothesis, P and F agree on the unit circle. Now by Theorem 14.1,

$$\sum_{x \in Z(F) \cap \mathbb{D}^2} \mu_F(x) = W(F|_{S^1}, 0) = W(P|_{S^1}, 0) = \sum_{x \in Z(P) \cap \mathbb{D}^2} \mu_P(x).$$

At a point $x \in Z(P)$, the multiplicity $\mu_P(x)$ is always $+1$, because the derivative $dP : \mathbb{C} \to \mathbb{C}$ is complex linear. If we view dP as a map from $\mathbb{R}^2$ to $\mathbb{R}^2$, then it preserves orientation, and so it has positive determinant. Then it follows that $|Z(F) \cap \mathbb{D}^2| \geq |Z(P) \cap \mathbb{D}^2|$. Since P and F agree outside of $\mathbb{D}^2$, we have $|Z(F)| \geq |Z(P)|$. □

There is a deeper theorem that generalizes this result to polynomials in many variables. To state this result, we first recall a little more about regular values.

PROPOSITION 14.3. (The Preimage Theorem on page 21 of [**GP**]) If $f : \mathbb{R}^m \to \mathbb{R}^n$ is a smooth map, and $y \in \mathbb{R}^n$ is a regular value, then $f^{-1}(y)$ is a smooth submanifold of dimension $m - n$.

The same definition and result make sense over the complex numbers. If $f : \mathbb{C}^m \to \mathbb{C}^n$ is a holomorphic map, and point $z \in \mathbb{C}^m$ is a regular point if $df_z : \mathbb{C}^m \to \mathbb{C}^n$ is surjective, and a point $w \in \mathbb{C}^n$ is a regular value if every point $z \in f^{-1}(w)$ is a regular point. If w is a regular value, then $f^{-1}(w)$ is a complex submanifold of $\mathbb{C}^m$ of (complex) dimension $m - n$.

Now we can state a generalization of Theorem 14.2 to polynomials of many variables. We identify $\mathbb{C}^n = \mathbb{R}^{2n}$ and use the Euclidean metric on $\mathbb{R}^{2n}$ to measure volumes.

THEOREM 14.4. (Federer) Suppose that $P : \mathbb{C}^n \to \mathbb{C}$ is a complex polynomial, and suppose that $F : \mathbb{R}^{2n} \to \mathbb{R}^2$ is a smooth function that agrees with P outside of the unit ball B^{2n}. Moreover, suppose that 0 is a regular value of P and F. Then

$$\text{Vol}_{2n-2}\, Z(P) \cap B^{2n} \leq Vol_{2n-2} Z(F) \cap B^{2n}. \tag{14.1}$$

This result says that the zero set of P does not waste any volume - it is as efficient as possible. This theorem implies that $Z(P)$ is a minimal surface. This result plays an important role in the theory of minimal surfaces and in differential geometry. Among other places, the proof appears in Federer's book on geometric measure theory [**Fed**].

The proof of this theorem is beyond the scope of this book, but we can make a couple comments about it. Suppose that L is a complex line in $\mathbb{C}^n$. For almost any such line, P does not vanish on $L \cap \partial B^{2n}$. In this case, Theorem 14.2 implies that $|Z(F) \cap L \cap B^{2n}| \geq |Z(P) \cap L \cap B^{2n}|$. In other words, for almost every complex line L, the intersection $L \cap (Z(F) \cap B^{2n})$ is bigger than $L \cap (Z(P) \cap B^{2n})$. From this information, we would like to conclude that $Z(F) \cap B^{2n}$ is bigger than $Z(P) \cap B^{2n}$. It is possible to do this using integral geometry.

Integral geometry studies the connection between the geometry of a surface $M \subset \mathbb{R}^n$ and the intersections $M \cap \pi$ for various planes $\pi \subset \mathbb{R}^n$. We will introduce integral geometry in Section 14.3 below, but we won't do enough to prove Theorem 14.4.

The proof of Theorem 14.4 is usually written in a different way using differential forms. It has had a significant influence in geometry - many other arguments modelled on it have appeared since then. This type of argument was dubbed a calibration argument by Harvey and Lawson who generalized it to many other settings. A good place to read about this material is their paper [**HL**].

A more recent result describes the topological efficiency of complex algebraic curves. It was conjectured by Milnor in the 60's and proven by Kronheimer and Mrowka in the 90's. We state a special case of their result here:

THEOREM 14.5. (Kronheimer-Mrowka [**KM**]) Suppose that $P : \mathbb{C}^2 \to \mathbb{C}$ is a complex polynomial, and suppose that $F : \mathbb{R}^4 \to \mathbb{R}^2$ is a smooth function that agrees with P outside of the unit ball B^4. Also, suppose that 0 is a regular value for P and F, and that $Z(P) \cap B^4$ and $Z(F) \cap B^4$ are connected. Then the genus of $Z(P) \cap B^4$ is at most the genus of $Z(F) \cap B^4$.

Since 0 is a regular value for P and F, $Z(P) \cap B^4$ and $Z(F) \cap B^4$ are both oriented surfaces. The theorem says that $Z(P) \cap B^4$ is topologically simpler than $Z(F) \cap B^4$ – the polynomial P does not waste any 'handles' of $Z(P)$. It seems to be unknown if this theorem has any generalizations to polynomials on $\mathbb{C}^n$ for $n > 2$.

14.2. The efficiency of real polynomials

The results in the last section describe beautiful special features of complex polynomials. Our next goal is to understand whether they have any analogue for real polynomials. At first, this seems impossible. By the Weierstrauss approximation theorem, any continuous function on a compact set in $\mathbb{R}^n$ can be approximated arbitrarily well by a real polynomial. Morally, a real polynomial can impersonate any function – conversely, nothing at all is special about real polynomials. It's easy to write down a real polynomial P and a competitor F which badly violate the volume inequality in Equation 14.1.

For example, consider the polynomial in one variable: $P(x) = x^2 - 1$ on the interval $[-2, 2]$. This polynomial has two zeroes on the interval $[-2, 2]$, but it is easy to find a smooth function which agrees with P outside of $[-2, 2]$ and which is strictly positive.

Interesting results appear when we shift our perspective from a single degree D polynomial to the whole space $\text{Poly}_D(\mathbb{R}^n)$. An individual polynomial $P \in \text{Poly}_D(\mathbb{R}^n)$ is not an efficient function, but the whole space $\text{Poly}_D(\mathbb{R}^n)$ is efficient compared to other spaces of functions with the same dimension. Moreover, the ideas in the proof are similar to the ideas in the proof of the finite field Nikodym theorem.

Let us set up some relevant terminology. If $\Omega \subset \mathbb{R}^n$ is an open set, and $V \subset C^0(\Omega, \mathbb{R})$ is a vector space of continuous functions from Ω to $\mathbb{R}$, then we define

$$\mathrm{Area}_\Omega V := \sup_{0 \neq f \in V} \mathrm{Vol}_{n-1} Z(f) \cap \Omega.$$

Now we can state a precise theorem about the efficiency of real polynomials.

THEOREM 14.6. (Gromov, [**Gr**]) Let n be any dimension and $D \geq 1$ be any degree. If $V \subset C^0(B^n, \mathbb{R})$ is a vector space of functions and $\mathrm{Dim}\, V = \mathrm{Dim}\, \mathrm{Poly}_D(\mathbb{R}^n)$, then

$$\mathrm{Area}_{B^n} V \geq c_n \mathrm{Area}_{B^n} \mathrm{Poly}_D(\mathbb{R}^n).$$

This theorem appears in Section 4.2 of [**Gr**], and there are related ideas in [**Gr2**]. The optimal constant c_n in this theorem is unknown.

The proof of the theorem is based on two steps. The first step is to show that $\mathrm{Area}\, \mathrm{Poly}_D(\mathbb{R}^n) \sim D$. This step is a classical result, which is based on the vanishing lemma.

PROPOSITION 14.7. (Crofton) There are constants $c_n < C_n$ so that $c_n D \leq \mathrm{Area}_{B^n} \mathrm{Poly}_D(\mathbb{R}^n) \leq C_n D$.

The next step is to show that $\mathrm{Area}_{B^n} V$ cannot be too small.

PROPOSITION 14.8. If $V \subset C^0(B^n, \mathbb{R})$ is a vector space of functions with $\mathrm{Dim}\, V \geq 2$, then $\mathrm{Area}_{B^n} V \geq c_n (\mathrm{Dim}\, V)^{1/n}$.

The proof of this proposition uses a version of the parameter counting argument. If $\mathrm{Dim}\, V$ is large, we have many parameters at our disposal, and we can tune the parameters to find a non-zero $f \in V$ which vanishes a lot. The argument uses the Stone-Tukey ham sandwich theorem.

These two propositions immediately imply Theorem 14.6. In particular, if we know that $\mathrm{Dim}\, V = \mathrm{Dim}\, \mathrm{Poly}_D(\mathbb{R}^n) \geq c_n D^n$, then we see that $\mathrm{Area}_{B^n} V \geq c_n D \geq c_n \mathrm{Area}_{B^n} \mathrm{Poly}_D(\mathbb{R}^n)$.

14.3. The Crofton formula in integral geometry

In this section, we sketch the proof of Proposition 14.7. The proof is based on integral geometry.

Suppose that P is a non-zero polynomial in $\mathrm{Poly}_D(\mathbb{R}^n)$. What is the maximal possible volume of $Z(P) \cap B^n$? If P is a product of D linear factors, then $Z(P)$ is a union of D planes. By choosing planes that go through the origin, we can arrange that $Z(P) \cap B^n$ has $(n-1)$-volume $\sim D$. We will prove that for any $P \in \mathrm{Poly}_D(\mathbb{R}^n)$, $Z(P) \cap B^{n-1}$ has volume $\lesssim D$.

In Exercise 2.3, we proved that if $P \in \mathrm{Poly}_D(\mathbb{F}_q^n)$, then $|Z(P)| \leq D q^{n-1}$. This estimate is sharp when $Z(P)$ consists of D parallel planes. The proof was based on the vanishing lemma, which says that if l is a line in $\mathbb{F}_q^n$, then either $Z(P) \cap l$ contains $\leq D$ points or else $l \subset Z(P)$. The lines contained in $Z(P)$ turn out to be rather rare. Just for heuristics, suppose that we knew that $|Z(P) \cap l| \leq D$ for every line l. Then it follows by averaging over all the lines that $|Z(P)|/|\mathbb{F}_q^n| \leq D/q$ giving $|Z(P)| \leq D q^{n-1}$. (In the real proof of Exercise 2.3, you have to be more careful to deal with the lines $l \subset Z(P)$.)

Our estimate for the volume of $Z(P) \subset \mathbb{R}^n$ follows similar ideas. If l is a line, then either $l \cap Z(P)$ contains $\leq D$ points, or else $l \subset Z(P)$. We will control the size

of $Z(P)$ by considering $|Z(P) \cap l|$ for each line $l \subset \mathbb{R}^n$ and averaging over all lines l. In the real case, we are averaging over an infinite set of lines. In the 1800's, Crofton figured out how to take such an average. Crofton found a formula to recover the volume of a hypersurface $S \subset \mathbb{R}^n$ in terms of the number of intersections in $S \cap l$ for all the lines $l \subset \mathbb{R}^n$. Crofton's formula says that the volume is an appropriate average of $|S \cap l|$ over all l.

Each line in $\mathbb{R}^n$ has a parametrization of the form $\gamma(t) = vt+b$, where $v \in S^{n-1}$ and b is perpendicular to v. Let $l(v,b)$ be the line parametrized by $\gamma(t) = vt + b$. (Each line now has exactly two parametrizations with opposite orientations, because $l(v,b) = l(-v,b)$.)

THEOREM 14.9. (Crofton) For each dimension $n \geq 2$ there is a constant $\alpha_n > 0$ so that the following holds. If S is a smooth hypersurface in $\mathbb{R}^n$, then

$$\mathrm{Vol}_{n-1} S = \alpha_n \int_{S^{n-1}} \left(\int_{v^\perp} |l(v,b) \cap S| db \right) dvol_{S^{n-1}}(v). \qquad (*)$$

We give some motivation for this theorem, but not a complete proof. The key point is that both sides of equation $(*)$ are invariant with respect to translating and rotating S. (The left hand side is clearly invariant with respect to translation and rotation. Checking that the right hand side is invariant with respect to rotation and translation is an exercise in multivariable calculus.)

By choosing α_n, we can arrange that formula $(*)$ holds when S is the unit $(n-1)$-cube $[0,1]^{n-1} \times \{0\}$. By invariance, $(*)$ also holds for any translation or rotation of the unit $(n-1)$-cube.

Next we notice that both sides of $(*)$ are additive with respect to disjoint unions. Let's define the right hand side of $(*)$ to be $Cr(S)$, the Crofton size of S. If S_1 and S_2 are disjoint hypersurfaces, then $\mathrm{Vol}\, S_1 \cup S_2 = \mathrm{Vol}\, S_1 + \mathrm{Vol}\, S_2$ and $Cr(S_1 \cup S_2) = Cr(S_1) + Cr(S_2)$.

Next, we subdivide the unit cube into A^{n-1} subcubes of side length A^{-1}, for some integer $A \geq 2$. By translation symmetry, they each have the same Crofton size. By additivity, we see that $Cr([0, A^{-1}]^{n-1} \times \{0\}) = A^{-(n-1)}$. In other words, the equation $(*)$ holds for cubes of side length A^{-1} for any integer $A \geq 1$.

Finally, we approximate an arbitrary smooth S by a finite union of small $(n-1)$-cubes. In fact, we let S_j be a sequence of approximations with finer cubes. For every j, we have $\mathrm{Vol}_{n-1} S_j = Cr(S_j)$. We only have to check that we can choose a sequence of approximations S_j so that $\mathrm{Vol}\, S_j \to \mathrm{Vol}\, S$ and $Cr(S_j) \to Cr(S)$. This last point is more technical, and we don't prove it here, but I hope that this discussion gives some hint why the Crofton formula should be true. For more information about the Crofton formula, see [**Sa**].

Assuming Crofton's formula, we can now bound the volumes of algebraic surfaces. Again we give the main idea but don't do all technical details. Let P be a non-zero polynomial of degree $\leq D$. Suppose first that $Z(P)$ is a smooth hypersurface. We know that $|l(v,b) \cap Z(P)| \leq D$ unless $l(v,b) \subset Z(P)$. Next we explain why the lines contained in $Z(P)$ contribute nothing to the integral in $Cr(Z(P))$. For each v, the set of $b \in v^\perp$ so that $l(v,b) \subset Z(P)$ has measure 0. Therefore, for each v, these lines contribute nothing to the inner integral.

Next we note that if $|b| > 1$, $l(v,b)$ does not intersect the unit ball B^n. In this case, $l(v,b) \cap (Z(P) \cap B^n)$ is empty. Combining the information we have gathered,

we see that

$$\mathrm{Vol}_{n-1} Z(P) \cap B^n \leq \alpha_n \int_{S^{n-1}} \left(\int_{b \in v^\perp, |b| \leq 1} D db \right) dvol_{S^{n-1}}(v).$$

For comparison, every line $l(v,b)$ with $|b| < 1$ intersects the unit sphere S^{n-1} in exactly two points. Therefore,

$$\mathrm{Vol}_{n-1} S^{n-1} = \alpha_n \int_{S^{n-1}} \left(\int_{b \in v^\perp, |b| \leq 1} 2 db \right) dvol_{S^{n-1}}(v).$$

Comparing the last two equations, we see that

$$Vol_{n-1} Z(P) \cap B^n \leq (D/2) \,\mathrm{Vol}_{n-1} S^{n-1}.$$

Incidentally, if D is even, the bound is exactly sharp, because $Z(P)$ could consist of $D/2$ spheres with center at 0 and radii arbitrarily close to 1.

To prove Proposition 14.7, one has to check that the same bound holds when $Z(P)$ has singular points. We don't address this technical point here.

Although there are some technical details to fill in, I would like to emphasize that morally the estimate for the volume of $Z(P) \cap B^n$ follows from the symmetry of $\mathbb{R}^n$ and the vanishing lemma. Our theorem says that a polynomial of degree $\leq D$ cannot vanish on a set of much larger volume than D planes, and so the theorem is a cousin of estimates for the number of zeroes of a polynomial $P : \mathbb{F}_q^n \to \mathbb{F}_q$ which we used in the proof of the finite field Nikodym theorem.

14.4. Finding functions with large zero sets

Let V be a vector space of functions in $C^0(B^n, \mathbb{R}^n)$. If the dimension of V is large, we want to prove that there exists a non-zero $F \in V$ so that $\mathrm{Vol}_{n-1} Z(F) \cap B^n$ is large. Morally, we are going to prove this by a parameter-counting argument. If $\mathrm{Dim}\, V$ is large, we have a lot of parameters to play with, and we will choose them so that F vanishes at a lot of places.

We will exploit the large dimension of V by using the Stone-Tukey ham sandwich theorem, which we discussed in Section 10.3.1.

THEOREM. (General ham sandwich theorem, Stone and Tukey, [**StTu**]) Let V be a vector space of continuous functions on $\mathbb{R}^n$. Let $U_1, ..., U_N \subset \mathbb{R}^n$ be finite volume open sets with $N < \mathrm{Dim}\, V$. For any function $f \in V \setminus \{0\}$, suppose that $Z(f)$ has Lebesgue measure 0. Then there exists a function $f \in V \setminus \{0\}$ which bisects each set U_i.

(Recall that a function f bisects a finite volume open set U if the sets $\{x \in U | f(x) > 0\}$ and $\{x \in U | f(x) < 0\}$ each have one half the volume of U.)

Using the ham sandwich theorem, we can now prove Proposition 14.8. Suppose that $V \subset C^0(B^n, \mathbb{R})$ is a vector space. If there is a non-zero $F \in V$ so that $Z(F)$ has positive Lebesgue measure, then $Z(F)$ has infinite $(n-1)$-dimensional volume, and so the conclusion holds. So we can assume that $Z(F)$ has zero Lebesgue measure for each $0 \neq F \in V$. Then we can apply the general ham sandwich theorem.

Let $N = \mathrm{Dim}\, V - 1$. Suppose that $U_1, ..., U_N$ are disjoint balls in B^n. We can choose $U_1, ..., U_N$ to have radius $\geq c_n N^{-1/n}$, for example by putting the centers of the balls on a cubical grid. Now we can choose a non-zero $F \in V$ that bisects each

U_i. This bisection forces $Z(F) \cap U_i$ to be fairly large for each i. We make this more precise as follows:

LEMMA 14.10. Suppose that F bisects the unit ball B^n. Then $Z(F) \cap B^n$ has $(n-1)$-volume at least $c_n > 0$.

We indicate a proof of this lemma in the exercises. Rescaling the ball, we see that $\mathrm{Vol}_{n-1}(Z(F) \cap U_i) \geq c_n N^{-\frac{n-1}{n}}$. Since the balls U_i are disjoint, it follows that the total volume $\mathrm{Vol}_{n-1}(Z(F) \cap B^n)$ is $\geq c_n N^{1/n}$. In other words, $\mathrm{Area}\, V \geq c_n N^{1/n} \geq c_n (\mathrm{Dim}\, V)^{1/n}$. This proves Proposition 14.8.

EXERCISE 14.1. We outline a proof of the bisection lemma, Lemma 14.10, based on ideas of Federer and Fleming from geometric measure theory. For simplicity, we explain the proof in the case that $Z(F)$ is a smooth $(n-1)$-dimensional submanifold of $\mathbb{R}^n$. But this argument can also be extended to more general situations.

Let $p \in B^n$, and let $\pi_p : \bar{B}^n \setminus \{p\} \to \partial B^n$ be the radial projection. In other words, if $x \in \bar{B}^n \setminus p$, then $\pi_p(x)$ is the point where the ray starting at p and passing through x intersects ∂B^n. (If $x \in \partial B^n$, then $\pi_p(x) = x$.) We will study how π_p behaves on $Z(F) \cap \bar{B}^n$ for various points $p \in B^n$.

If $F(p) > 0$, then the image $\pi_p(Z(F) \cap \bar{B}^n)$ contains all the points in ∂B^n where $F \leq 0$. To see this, suppose that $F(p) > 0$ and suppose that $y \in \partial B^n$ with $F(y) \leq 0$. Then there must be some point x on the segment from p to y where $F(x) = 0$, and $\pi_p(x) = y$.

After possibly replacing F by $-F$, we can assume that

$$\mathrm{Vol}_{n-1}(\{y \in \partial B^n | F(y) \leq 0\}) \geq (1/2)\, \mathrm{Vol}_{n-1}\, \partial B^n.$$

Therefore, we see that for all p with $F(p) > 0$,

$$\mathrm{Vol}_{n-1}(\pi_p(Z(F) \cap \bar{B}^n) \gtrsim 1.$$

This suggests the following question: If $S \subset \bar{B}^n$ is an $(n-1)$-dimensional manifold, is it true that $\mathrm{Vol}_{n-1}(\pi_p(S)) \leq C_n\, \mathrm{Vol}_{n-1}(S)$? In general the answer to this question is no. For instance, if S is a small sphere centered at p, then $\mathrm{Vol}_{n-1}(S)$ is small, but $\pi_p(S) = \partial B^n$. However, for a fixed surface S, Federer and Fleming proved that $\mathrm{Vol}_{n-1}(\pi_p(S)) \leq C_n\, \mathrm{Vol}_{n-1}(S)$ for many points $p \in B^n$. In particular, Federer and Fleming proved that, for any $\rho < 1$, for any $(n-1)$-dimensional submanifold S in B^n,

$$\int_{B_\rho} \mathrm{Vol}_{n-1}(\pi_p(S)) dp \leq C(\rho, n)\, \mathrm{Vol}_{n-1}(S). \qquad (*)$$

Use $(*)$ to prove the bisection lemma.

Next we outline the proof of $(*)$. Suppose that $S \subset B^n$ is an $(n-1)$-dimensional submanifold and that $\rho < 1$. First, using multivariable calculus, prove that

$$\mathrm{Vol}_{n-1}\, \pi_p(S) \leq C(\rho, n) \int_S |p - x|^{-(n-1)}\, \mathrm{dvol}_S(x). \qquad (14.2)$$

Next plug this formula into $\int_{B(\rho)} \mathrm{Vol}_{n-1}\, \pi_p(S) dp$ and use Fubini.

14.5. An application of the polynomial method in geometry

We now give an application of the polynomial method to a geometry problem that does not mention polynomials. The problem has to do with area-expanding embeddings.

If $\Omega \subset \mathbb{R}^3$ is an open set, recall that an embedding $\phi : \Omega \to \mathbb{R}^3$ is a smooth map so that there is a smooth inverse map $\phi^{-1} : \phi(\Omega) \to \Omega$. In particular, if ϕ is an embedding then ϕ is injective and $d\phi_x$ is an isomorphism for every $x \in \Omega$. An embedding ϕ is called area-expanding if it does not decrease the area of any 2-dimensional surface in Ω. In other words, for any 2-dimensional submanifold $\Sigma \subset \Omega$, $\mathrm{Vol}_2\, \phi(\Sigma) \geq \mathrm{Vol}_2\, \Sigma$.

For example, the linear map $\phi_\epsilon(x, y, z) = (\epsilon x, \epsilon^{-1} y, \epsilon^{-1} z)$ is an area-expanding embedding for any ϵ in the range $0 < \epsilon < 1$. If Σ is a square in the xy-plane, then $\mathrm{Area}\, \phi_\epsilon(\Sigma) = \mathrm{Area}\, \Sigma$. The same holds for a square in the xz-plane, and a square in the yz-plane gets much bigger under ϕ_ε. With a little more work, the reader can check that the same holds for a small square at any angle, which implies that ϕ_ϵ is area-expanding. (For a further introduction to area-expanding embeddings, see [**Gu1**].) The map ϕ_ε sends the unit cube to a rectangular solid with one small axis and two large axes, something like a long thin sheet.

There are also lots of non-linear area-expanding embeddings. To visualize one example, let us imagine that the domain is made out of rubber. Anything you can do with the rubber without stretching/contracting it much is approximately an area-expanding embedding. For example, if the domain is a long thin rubber sheet, you could roll it up like a carpet. If we begin with an embedding that doesn't stretch or contract lengths by more than 20 %, then we can make an honest area-expanding embedding by composing it with the map $x \mapsto 2x$.

We have now met two area-exanding embeddings: the linear map ϕ_ε which distorts lengths a lot, and the carpet-rolling map which approximately preserves lengths but is very non-linear. To construct complicated examples, notice that the composition of two area-expanding embeddings is an area-expanding embedding. For example, we can first send the unit cube to a thin square sheet and then roll up the sheet into a tube. Then we could use another linear map to flatten the tube into a thin sheet, and then we could roll up that thin sheet...

Now we come to a question about area-expanding embeddings. Given two rectangular solids Ω_1 and Ω_2, when is there an area-expanding embedding from Ω_1 into Ω_2? Here is an interesting subcase:

Suppose that T is a rectangular solid of dimensions $1 \times 1 \times L$, for some $L \geq 1$. The letter T stands for tube. Suppose that P is a rectangular solid of dimensions $\epsilon \times S \times S$, where $\epsilon < 1 < S$. The letter P stands for pancake. For which values of L, ϵ, S can we find an area expanding embedding from T into P? First we sketch a construction.

Proposition 14.11. If $L < (1/10)\epsilon^2 S^2$, then there is an area-expanding embedding $T \to P$.

We sketch the proof. As we saw above, there is an area-expanding embedding $\phi_\epsilon : T \to [0, \epsilon] \times [0, \epsilon^{-1}] \times [0, \epsilon^{-1} L]$. Next we want to find an area-expanding embedding from this rectangle into P. Notice that both this rectangle and P have first dimension ϵ. We will look for a map of the form $(x, y, z) \to (x, \psi(y, z))$, where ψ is a length-expanding embedding from $[0, \epsilon^{-1}] \times [0, \epsilon^{-1} L]$ into $[0, S]^2$. It just remains to find such a ψ.

The inequality $L < (1/10)\epsilon^2 S^2$ is equivalent to

$$\mathrm{Area}\left([0, \epsilon]^{-1} \times [0, \epsilon^{-1} L]\right) \leq (1/10)\, \mathrm{Area}([0, S]^2).$$

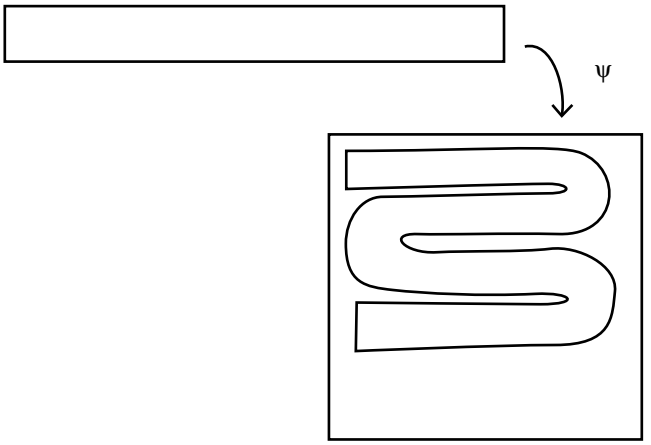

FIGURE 14.1. Folding up a rectangle inside a square.

Since $L \geq 1$, it also implies that $\epsilon^{-1} \leq (1/3)S$. It's now relatively easy to fold up the rectangle $[0, \epsilon^{-1}] \times [0, \epsilon^{-1}L]$ inside of $[0, S]^2$ without shrinking any lengths. Figure 14.1 shows a picture of how the image of such a map may look.

This finishes our sketch of the proof.

The construction we just described only works if L is smaller than $\epsilon^2 S^2$. The question arises whether there is a better construction that would work even if L is much larger than $\epsilon^2 S^2$. There are lots of highly non-linear area expanding embeddings, and so there are lots of possible constructions. But it turns out that our construction is optimal up to a constant factor.

THEOREM 14.12. There is a constant C so that the following holds. If $\phi : T \to P$ is an area-expanding embedding, then $L < C\epsilon^2 S^2$.

The statement of this theorem has nothing to do with polynomials, but we will prove it using polynomials. The proof is somewhat analogous to the proof of the finite field Nikodym theorem.

PROOF. Let $V \subset C^0(P, \mathbb{R})$ be the vector space of polynomials in the variables y, z with degree $\leq L^{1/2}$. We have $\operatorname{Dim} V \sim L$.

We evaluate $\operatorname{Area}_P V$. If Q is a polynomial in y, z of degree $\leq L^{1/2}$, then the Crofton formula implies that the length of $Z(Q) \cap [0, S]^2$ is $\lesssim L^{1/2} S$. Now if we think of Q as a map from P to $\mathbb{R}$ which doesn't depend on the x coordinate, then the area of $Z(Q) \cap P$ is $\lesssim \epsilon L^{1/2} S$.

$$\operatorname{Area}_P V \lesssim L^{1/2} \epsilon S. \tag{1}$$

Consider the pullback $\phi^* V \subset C^0(T, \mathbb{R})$. (If $F : P \to \mathbb{R}$, then recall that the pullback $\phi^* F : T \to \mathbb{R}$ is just defined by $\phi^* F(x) = F(\phi(x))$. The pullback $\phi^* V$ is just the set of all $\phi^* F, F \in V$. It's easy to check that $\phi^*(F_1 + F_2) = \phi_* F_1 + \phi^* F_2$ and $\phi^*(\lambda F) = \lambda \phi^* F$ for $\lambda \in \mathbb{R}$. Therefore, $\phi^* V$ is a vector space with the same dimension as V.)

Because ϕ is area-expanding, we claim that

$$\operatorname{Area}_T \phi^* V \leq \operatorname{Area}_P V. \tag{2}$$

To see this, consider any function $F \in V$. Since $\phi : T \to P$ is area-expanding, we have

$$\text{Area}\, Z(\phi^* F) \cap T \leq \text{Area}\, Z(F) \cap \phi(T) \leq \text{Area}\, Z(F) \cap P \leq \text{Area}_P\, V.$$

Next we estimate $\text{Area}_T\, \phi^* V$ using the Stone-Tukey ham sandwich theorem. We have $\text{Dim}\, \phi^* V \sim L$. The tube T contains $\sim L$ disjoint balls of radius 1. By the ham sandwich theorem, we can choose a non-zero $G \in \phi^* V$ which bisects $\sim L$ disjoint unit balls in T. Therefore, $\text{Area}\, Z(G) \cap T \gtrsim L$. Therefore,

$$L \lesssim \text{Area}_T\, \phi^* V. \tag{3}$$

Assembling equations 3, 2, and 1, we get

$$L \lesssim \text{Area}_T\, \phi^* V \leq \text{Area}_P\, V \lesssim L^{1/2} S \epsilon.$$

Rearranging gives $L \lesssim S^2 \epsilon^2$. □

Why are polynomials the right space of functions to use in this argument? They work because polynomials are the most efficient space of functions. We chose a space V with dimension $\sim L$. It was crucial that $\text{Area}_P\, V$ was approximately minimal among all spaces of dimension $\sim L$. As we saw in the last section, polynomials have approximately minimal area. Any space with approximately minimal area would work just as well. (But can you find any other such space... ?)

This proof is analogous to the proof of the finite field Nikodym theorem, and we end this chapter by comparing the two proofs. They are both proofs by contradiction.

For the finite field Nikodym theorem, we suppose that we have a small Nikodym set $N \subset \mathbb{F}_q^n$. By parameter counting, there must be a polynomial $Q(x_1, ..., x_n)$ of small degree that vanishes on N. By the vanishing lemma, Q vanishes at every point of $\mathbb{F}_q^n$. Now the polynomial Q vanishes too much, giving a contradiction.

For Theorem 14.12, we suppose that we have an area-expanding embedding $\phi : T \to P$ with dimensions obeying $L \gg \varepsilon^2 S^2$. By parameter counting, there must be a polynomial $Q(y, z)$ of controlled degree so that $\phi^* Q$ bisects a lot of unit cubes in T. Since ϕ is area-expanding, $Q(y, z)$ must vanish on a surface of large area in P. Comparing with the Crofton formula, we see that Q vanishes too much, giving a contradiction.

Exercise 14.2. (**) To end this section, we mention a difficult open problem. Theorem 14.6 says that $\text{Area}_{B^n}\, \text{Poly}_D(\mathbb{R}^n) \leq C_n\, \text{Area}_{B^n}\, V$, whenever $\text{Dim}\, V = \text{Dim}\, \text{Poly}_D(\mathbb{R}^n)$, for some constant C_n. It would be very interesting to prove a theorem in this spirit with constant $C_n = 1$. I believe that the sharp constant in Theorem 14.6 is bigger than 1, but there is an analogous question on real projective space where I think the sharp constant may be 1. This question requires a little more background in differential geometry or algebraic geometry to state.

Let $\text{Poly}_{=D}(\mathbb{R}^{n+1})$ be the space of homogeneous polynomials on $\mathbb{R}^{n+1}$ of degree exactly D. For any $P \in \text{Poly}_{=D}$, we can define $Z(P) \subset \mathbb{RP}^n$. Now P is not a function on $\mathbb{RP}^n$. Instead P is a section of a line bundle over $\mathbb{RP}^n$, called $O(D)$. So

we can think of $\mathrm{Poly}_{=D}(\mathbb{R}^{n+1})$ as a vector space of sections of $O(D)$. Suppose that V is another vector space of sections of $O(D)$, with $\mathrm{Dim}\, V = \mathrm{Dim}\, \mathrm{Poly}_{=D}(\mathbb{R}^{n+1})$. Does it follow that

$$\mathrm{Area}_{\mathbb{RP}^n}\, \mathrm{Poly}_{=D+1}(\mathbb{R}^{n+1}) \leq \mathrm{Area}_{\mathbb{RP}^n}\, V?$$

If it is true, this result would be a sharp statement about the efficiency of real polynomials.

https://doi.org/10.1090//ulect/064/15

CHAPTER 15

Harmonic analysis and the Kakeya problem

At the beginning of the book, in Chapter 2, we discussed the proof of the finite field Kakeya conjecture. This proof started many of the investigations we have described in the book. The finite field Kakeya problem is a toy problem for the original Kakeya problem in Euclidean space, an important open problem in harmonic analysis. In this chapter, we explain the Kakeya problem and its connections to harmonic analysis.

We also discuss how much the polynomial method has been able to say about the original Kakeya problem and about harmonic analysis. We explain some of the difficulties in trying to adapt the proof of finite field Kakeya to the Euclidean case, and we give a succesful application of the polynomial method to an easier problem related to Kakeya.

More broadly, this chapter is about some interactions between combinatorics, geometry, and analysis in Euclidean space $\mathbb{R}^n$. We will build up to the Kakeya problem, beginning with some simpler and more classical connections between combinatorics and geometry.

15.1. Geometry of projections and the Sobolev inequality

15.1.1. Loomis-Whitney Inequality. The Loomis-Whitney inequality is a geometric inequality related to the isoperimetric inequality. From the combinatorial point of view, it is a special case of the joints problem when the lines are axis-parallel.

Let X be a set of unit cubes in the unit cubical lattice in $\mathbb{R}^n$, and let $|X|$ be its volume. Let π_j be the projection onto the coordinate hyperplane perpendicular to the x_j-axis. If $\pi_j(X)$ is small for all j, what can we say about $|X|$?

THEOREM 15.1. ([**LW**]) If $|\pi_j(X)| \leq A$, then $|X| \lesssim A^{\frac{n}{n-1}}$.

Loomis and Whitney proved an estimate with a sharp constant: $|X| \leq A^{\frac{n}{n-1}}$. Their original proof uses Holder's inequality repeatedly.

This inequality is one of my favorite problems to give students. We can describe it informally in the following way. We have a set X which appears small when viewed from any coordinate direction. Does that mean that the set is actually small? Even more informally, does a large object always look large?

I have spent some time trying to find the most direct and least computational proof that I can. Here we give a proof which uses some of the inductive structure from the proof of the joints theorem in Section 2.5. It doesn't give the sharp constant, but the computations are fairly simple.

Define a *column* to be the set of cubes obtained by starting at any cube and taking all cubes along a line in the x_j-direction, for some j.

LEMMA 15.2. If $|\pi_j(X)| \leq B$ for every j, then there exists a column of cubes with between 1 and $B^{\frac{1}{n-1}}$ cubes of X.

PROOF. Suppose not, so every column has $> B^{\frac{1}{n-1}}$ cubes. This means that there are $> B^{\frac{1}{n-1}}$ cubes of X on some line parallel to the x_1-axis. Call this line A_1. If p lies in A_1 and lies in a cube of X, then the line through p parallel to the x_2-axis must intersect $> B^{\frac{1}{n-1}}$ cubes of X. Let A_2 be the 2-plane containing A_1 and parallel to the (x_1, x_2)-plane. It must intersect $> B^{\frac{2}{n-1}}$ cubes of X. Proceeding this way, we find an $(n-1)$-plane A_{n-1}, parallel to the $(x_1, ..., x_{n-1})$-plane, intersecting $> B$ cubes of X. But then $|\pi_n(X)| > B$, giving a contradiction. □

The Loomis-Whitney inequality follows from this Lemma by induction.

COROLLARY 15.3. If $\sum_j |\pi_j(X)| \leq B$, then $|X| \leq \sum_{b=1}^{B} b^{\frac{1}{n-1}}$.
Therefore, $|X| \leq B^{\frac{n}{n-1}}$.

This corollary implies that if $|\pi_j(X)| \leq A$ for each j, then $|X| \leq (nA)^{\frac{n}{n-1}}$, proving Theorem 15.1.

PROOF OF COROLLARY 15.3. We proceed by induction on B. The case $B = 1$ is trivial.

Let X' be X with its smallest column removed. Removing a column reduces the size of $\pi_j(X)$ for some j, and so $\sum_j |\pi_j(X')| \leq B - 1$. By induction on B, we can assume that

$$|X'| \leq \sum_{b=1}^{B-1} b^{\frac{1}{n-1}}.$$

By Lemma 15.2, the column that we removed contained at most $B^{\frac{1}{n-1}}$ cubes. Therefore,

$$|X| \leq |X'| + B^{\frac{1}{n-1}} \leq \left(\sum_{b=1}^{B-1} b^{\frac{1}{n-1}} \right) + B^{\frac{1}{n-1}}.$$

□

We now generalize the Loomis-Whitney inequality to open sets instead of just unions of unit cubes.

THEOREM 15.4 (more general Loomis-Whitney). If U is an open set in $\mathbb{R}^n$ with $|\pi_j(U)| \leq A$, then $|U| \lesssim A^{\frac{n}{n-1}}$.

PROOF. Take $U_\varepsilon \subset U$ be a union of ε-cubes in ε-lattice. Then $|U_\varepsilon| \lesssim A^{\frac{n}{n-1}}$ and $|U_\varepsilon| \to |U|$. □

As a corollary, we can prove a version of the isoperimetric inequality with a non-sharp constant.

COROLLARY 15.5 (Isoperimetric inequality). If U is a bounded open set in $\mathbb{R}^n$, then

$$\mathrm{Vol}_n(U) \lesssim \mathrm{Vol}_{n-1}(\partial U)^{\frac{n}{n-1}}.$$

PROOF. Since U is bounded, any line that intersects U must also intersect ∂U. Therefore, $\pi_j(U) \subset \pi_j(\partial U)$. The projection map π_j can only decrease $(n-1)$-dimensional volumes and so we get the inequality $|\pi_j(U)| \leq \mathrm{Vol}_{n-1}(\partial U)$. Now applying Theorem 15.4, we get the bound

$$|U| \lesssim (\max_j |\pi_j(U)|)^{\frac{n}{n-1}} \leq \mathrm{Vol}_{n-1}(\partial U)^{\frac{n}{n-1}}.$$

□

15.1.2. Sobolev Inequality. The Loomis-Whitney inequality is a combinatorial/geometric inequality about Euclidean space. Now we turn to its consequences in analysis. We will use Theorem 15.4 to prove the Sobolev inequality, a fundamental estimate in analysis. The Sobolev inequality relates the size of a function u and the size of its gradient ∇u.

We write $C^1_{\mathrm{comp}}(\mathbb{R}^n)$ for the space of C^1 compactly supported functions on $\mathbb{R}^n$. Suppose that $u \in C^1_{\mathrm{comp}}(\mathbb{R}^n)$ satisfies $\int |\nabla u| = 1$. How big can the function u be? Perhaps surprisingly, if $n \geq 2$, then $\max |u|$ may be arbitrarily large. (We describe an example in the exercises below.) Nevertheless, the restriction $\int |\nabla u| = 1$ does the limit the size of u for a well-chosen notion of size.

Recall that the L^p-norm $||u||_{L^p}$ is given by

$$||u||_{L^p} = \left(\int |u|^p \right)^{1/p}$$

For $p = \frac{n}{n-1}$, the L^p norm of $u \in C^1_{\mathrm{comp}}(\mathbb{R}^n)$ is controlled by $\int |\nabla u|$.

THEOREM 15.6 (Sobolev inequality). If $u \in C^1_{\mathrm{comp}}(\mathbb{R}^n)$, then

$$||u||_{L^{\frac{n}{n-1}}} \lesssim ||\nabla u||_{L^1}.$$

The Sobolev inequality is true only for this value of p. (Again see the exercises.)

To understand the size of u, the following definition is useful.

$$S_u(h) := \{x \in \mathbb{R}^n \text{ so that } |u(x)| > h\}.$$

Knowing the volume of $S_u(h)$ for different h encodes a lot of information about the size of u (including the L^p norms of u). The norm $\|\nabla u\|_{L^1}$ controls the size of the projection of $S_u(h)$ by the following lemma.

LEMMA 15.7. If $u \in C^1_{\mathrm{comp}}(\mathbb{R}^n)$, then for any j,

$$|\pi_j(S_u(h))| \leq h^{-1} \cdot ||\nabla u||_{L^1}.$$

PROOF. For $x \in S_u(h)$, take a line ℓ in the x_j-direction. It eventually reaches a point x' where $u(x') = 0$, so $\int_\ell |\nabla u| \geq h$ by the fundamental theorem of calculus. This means that

$$||\nabla u||_{L^1} \geq \int_{\pi_j(S_u(h)) \times \mathbb{R}} |\nabla u| = \int_{\pi_j(S_u(h))} \left(\int_{\mathbb{R}} |\nabla u| dx_j \right) dx_{\mathrm{other}} \geq |\pi_j(S_u(h))| \cdot h.$$

□

If we combine the Loomis-Whitney theorem, Theorem 15.4, with Lemma 15.7, we get the following estimate:

$$|S_u(h)| \lesssim h^{-\frac{n}{n-1}} \cdot ||\nabla u||_{L^1}^{\frac{n}{n-1}}. \tag{15.1}$$

This inequality is closely related to the Sobolev inequality, but it's a little weaker. To relate $\|u\|_{L^p}$ with the sizes $S_u(h)$, we can proceed as follows. For each integer k, define $T_u(k)$ to be the set where $|u|$ is roughly 2^k. More precisely:

$$T_u(k) := \{x \in \mathbb{R}^n \text{ so that } 2^k < |u(x)| \leq 2^{k+1}\}.$$

Then for any p,

$$\|u\|_{L^p}^p = \int |u|^p \sim \sum_{k \in \mathbb{Z}} |T_u(k)| 2^{kp}.$$

On the other hand, $|T_u(k)| \leq |S_u(2^k)|$. By Equation 15.1, we get the estimate

$$|T_u(k)| 2^{k \frac{n}{n-1}} \lesssim \|\nabla u\|_{L^1}^{\frac{n}{n-1}}.$$

We don't get any bound for $\|u\|_{L^{\frac{n}{n-1}}}^{\frac{n}{n-1}} \sim \sum_{k \in \mathbb{Z}} |T_u(k)| 2^{k \frac{n}{n-1}}$, but we get a good bound for each term in the sum. Equation 15.1 is a weaker version of the Sobolev inequality. To get the full Sobolev inequality, we have to be a little more careful about how the different scales relate to each other.

Proof of Theorem 15.6.

Lemma 15.8 (More careful version of Lemma 15.7). If $u \in C^1_{\text{comp}}(\mathbb{R}^n)$, then for any j, we have the inequality

$$|\pi_j T_u(k)| \lesssim 2^{-k} \int_{T_u(k-1)} |\nabla u|.$$

The new detail here is that on the right-hand side, instead of taking the integral of $|\nabla u|$ over all of $\mathbb{R}^n$, we only have to integrate $|\nabla u|$ over $T_u(k-1)$.

Proof. Suppose $x \in T_u(k)$. In other words, $2^k < |u(x)| \leq 2^{k+1}$.

Consider a line ℓ_x in the x_j-direction through x. There is a point x' on ℓ with $u(x') = 0$. Between x and x', there is a region of ℓ_x where $|u|$ is between 2^{k-1} and 2^k – this region lies in $T_u(k-1)$. In this region, there must be at least one line segment where $|u| = 2^k$ at one endpoint and 2^{k-1} at the other endpoint. By the fundamental theorem of calculus we see that

$$\int_{\ell_x \cap T_u(k-1)} |\nabla u| \geq \frac{1}{2} 2^k.$$

We note that ℓ_x does not depend on the x_j coordinate of x - only on the other coordinates. Now we proceed as in the proof of Lemma 15.7:

$$\int_{T_u(k-1)} |\nabla u| \geq \int_{\pi_j(T_u(k))} \left(\int_{\ell_x \cap T_u(k-1)} |\nabla u| dx_j \right) dx_{other} \geq \frac{1}{2} 2^k |\pi_j(T_u(k))|.$$

□

If we combine the Loomis-Whitney theorem, Theorem 15.4, with Lemma 15.8, we get the following slightly stronger version of Inequality 15.1:

$$|T_u(k)| \lesssim 2^{-k \frac{n}{n-1}} \left(\int_{T_u(k-1)} |\nabla u| \right)^{\frac{n}{n-1}}. \tag{15.2}$$

We can now finish the proof of the Sobolev inequality by summing the contribution from different values of k.

$$\int |u|^{\frac{n}{n-1}} \sim \sum_{k=-\infty}^{\infty} |T_u(k)| 2^{k\frac{n}{n-1}} \lesssim \sum_k \left(\int_{T_u(k-1)} |\nabla u| \right)^{\frac{n}{n-1}} \leq \left(\int_{\mathbb{R}^n} |\nabla u| \right)^{\frac{n}{n-1}},$$

where in the last step we move the sum inside the $\frac{n}{n-1}$-power. □

EXERCISE 15.1. Examples for the Sobolev inequality. Let u be any function in $C^1_{comp}(\mathbb{R}^n)$. Define u_λ by rescaling the space variable:

$$u_\lambda(x) := u(x/\lambda).$$

Consider functions of the form hu_λ. If $p \neq \frac{n}{n-1}$, show that we can choose a sequence h_j, λ_j so that $\|u\|_{L^p} \to \infty$ and $\|\nabla u\|_{L^1} \to 0$.

15.2. L^p estimates for linear operators

Suppose that T is an operator that takes functions on $\mathbb{R}^n$ to functions on $\mathbb{R}^n$. Harmonic analysts often study L^p estimates for the operator. In other words, they try to find all pairs of exponents p, q so that, for some constant C,

$$\|Tf\|_{L^q(\mathbb{R}^n)} \leq C \|f\|_{L^p(\mathbb{R}^n)}.$$

For the reader who hasn't studied L^p estimates before, let us briefly try to explain the kind of information contained in these inequalities. Suppose that the input function f has support of volume V and suppose that on the support $h/2 \leq |f| \leq h$. We want to understand the size of the output function Tf. Recall that $S_{Tf}(H)$ is the set of x where $|Tf(x)| > H$. In terms of V, h, we would like to understand how big $|S_{Tf}(H)|$ can be. This type of estimate is closely related to L^p estimates.

PROPOSITION 15.9. Suppose that T obeys the inequality $\|Tf\|_{L^q(\mathbb{R}^n)} \leq C\|f\|_{L^p(\mathbb{R}^n)}$. If the measure of the support of f is equal to V, and if $|f| \leq h$ everywhere, then

$$|S_{Tf}(H)| \leq C^q V^{q/p} (h/H)^q.$$

PROOF. We have

$$|S_{Tf}(H)|^{1/q} H \leq \|Tf\|_{L^q} \leq C\|f\|_{L^p} \leq CV^{1/p} h.$$

Rearranging, we get

$$|S_{Tf}(H)| \leq C^q V^{q/p} (h/H)^q.$$

□

In summary, understanding L^p estimates for an operator T describes to what extent T can take a short wide input f and produce a tall thin output, and to what extent it can take a tall thin input and produce a short wide output.

The linear operators that we study in this chapter will be convolutions. We do a quick review of convolutions.

If f, g are functions from $\mathbb{R}^n$ to $\mathbb{R}$ (or to $\mathbb{C}$), we define the *convolution* to be

$$(f * g)(x) := \int_{\mathbb{R}^n} f(y) g(x-y) dy = \int_{\mathbb{R}^n} f(x-y) g(y) dy.$$

If you haven't seen this definition before, it might help to explain it with the following story. Suppose there is a factory at 0 which generates a cloud of pollution

centered at 0 described by $g(-y)$. If the density of factories at x is $f(x)$, then the final observed pollution is $f * g$.

The first operator we will study is convolution by $|x|^{-\alpha}$, for $0 < \alpha < n$. We define

$$T_\alpha f := f * |x|^{-\alpha}.$$

In other words,

$$T_\alpha f(x) := \int f(y)|x-y|^{-\alpha} dy.$$

We will take α in the range $0 < \alpha < n$, and we will assume (at least initially) that f is a bounded measurable function with compact support. These two assumptions guarantee that integral above converges for each x, and that $T_\alpha f$ is continuous.

To get a feel for this operator, we consider the case when f is the characteristic function of a ball. We write B_r for the ball of radius r centered at 0, and χ_{B_r} for the characteristic function of B_r. For these examples, we estimate $T\chi_{B_r}$.

$$|T_\alpha \chi_{B_r}(x)| \sim \begin{cases} r^n \cdot r^{-\alpha} & \text{if } |x| \leq r \\ r^n \cdot |x|^{-\alpha} & \text{if } |x| > r. \end{cases}$$

Suppose that $A \subset \mathbb{R}^n$ has volume r^n. It may seem intuitive that for any height H,

$$|S_{T\chi_A}(H)| \lesssim |S_{T\chi_{B_r}}(H)|.$$

This inequality turns out to be true. It follows from L^p estimates for the operator T_α which were first proven by Hardy and Littlewood and (independently) by Sobolev.

THEOREM 15.10. (Hardy-Littlewood-Sobolev) If $p > 1$ and $\alpha = n(1 - \frac{1}{q} + \frac{1}{p})$, then $\|T_\alpha f\|_q \leq C(n,p,q)\|f\|_p$.

This inequality has many applications in analysis and partial differential equations. We sketch one in the exercises at the end of this section. The algebraic restriction on p, q, n, α looks a little complicated, but it has a simple interpretation: these are just the exponents so that the inequality $\|T_\alpha f\|_q \lesssim \|f\|_p$ holds in the simple example $f = \chi_{B_r}$. More precisely, we can say that

PROPOSITION 15.11. Fix a dimension n and consider the linear operator T_α. The following are equivalent:

(1) There exists a constant C so that for every $r > 0$, $\|T_\alpha \chi_{B_r}\|_q \leq C\|\chi_{B_r}\|_p$.
(2) $p > 1$ and $\alpha = n(1 - \frac{1}{q} + \frac{1}{p})$.

(We leave the proof as an exercise for the reader.)

The proof of Theorem 15.10 is based on geometric/combinatorial estimates about the intersection patterns of balls in $\mathbb{R}^n$. It is our second example of how geometry, combinatorics, and analysis interact. We give the proof in the next section.

EXERCISE 15.2. Prove Proposition 15.11.

EXERCISE 15.3. Assuming Theorem 15.10, prove the following. Suppose that $A \subset \mathbb{R}^n$ has volume r^n. For any r and any H,

$$|S_{T\chi_A}(H)| \leq C(n,\alpha)|S_{T\chi_{B_r}}(H)|.$$

EXERCISE 15.4. If $u \in C^1_{comp}(\mathbb{R}^n)$, prove that at each point x,

$$|u(x)| \lesssim T_{n-1}|\nabla u|.$$

Applying Theorem 15.10, prove the following version of the Sobolev inequality. Suppose that $u \in C^1_{comp}(\mathbb{R}^n)$, and $1 < p < n$ and $n - 1 = n(1 - \frac{1}{q} + \frac{1}{p})$. Then

$$\|u\|_{L^q(\mathbb{R}^n)} \leq C(n,p)\|\nabla u\|_{L^p(\mathbb{R}^n)}.$$

15.3. Intersection patterns of balls in Euclidean space

Our next topic in the geometry of Euclidean space is to study the combinatorics of how balls overlap. Here is a typical question, called the ball doubling problem. Suppose that B_i is a finite list of balls in $\mathbb{R}^n$. Let $2B_i$ be the ball with the same center as B_i and twice the radius. Is there a universal constant C_n so that for any finite set of balls in $\mathbb{R}^n$,

$$|\cup_i 2B_i| \leq C_n|\cup_i B_i|?$$

If we take a single ball B, then $|2B| = 2^n|B|$. It looks plausible that for any collection of balls B_i, $|\cup_i 2B_i| \leq 2^n|\cup_i B_i|$, but it is not obvious how to prove even with a larger constant C_n. We will prove such an estimate using the Vitali covering lemma, a fundamental result about balls in Euclidean space.

LEMMA 15.12. (Vitali Covering Lemma) If $\{B_i\}_{i\in I}$ is a finite collection of balls in $\mathbb{R}^n$, then there exists a subcollection $J \subset I$ such that $\{B_j\}_{j\in J}$ are disjoint but $\bigcup_{i\in I} B_i \subset \bigcup_{j\in J} 3B_j$.

PROOF. Let B_{j_1} be a ball with maximal radius, and add j_1 into J. Let B_{j_2} be a ball disjoint from B_{j_1}, with maximal radius among all the choices. Add j_2 to J. We continue in this way, adding a ball of maximal radius disjoint from all the balls in J, until no more balls are available. Suppose that $i \notin J$. We have to show that $B_i \subset \cup_{j\in J} 3B_j$. Since i was not added to J, there must be a first j_k so that $B_i \cap B_{j_k}$ is non-empty. Since B_i is disjoint from all the previous balls of J, the radius of B_{j_k} must be at least the radius of B_i. But then $B_i \subset 3B_{j_k}$. □

Remark: This Lemma actually holds in any metric space.

From the Vitali covering lemma, we get the following estimate for the ball doubling problem.

LEMMA 15.13. (Ball doubling) If $\{B_i\}_{i\in I}$ is a finite collection of balls, then $|\bigcup 2B_i| \leq 6^n|\bigcup B_i|$.

PROOF. We apply the Vitali covering lemma to the set of balls $2B_i$. By the Vitali Covering Lemma, there exists a subcollection $J \subset I$ such that $\{2B_j\}_{j\in J}$ are disjoint but $\bigcup_{i\in I} 2B_i \subset \bigcup_{j\in J} 6B_j$. In particular, $\{B_j\}_{j\in J}$ are disjoint. Hence $|\bigcup 2B_i| \leq |\bigcup 6B_j| \leq 6^n \sum_j |B_j| = 6^n|\bigcup B_j|$. □

(The reader may be interested to know whether the sharp constant is 2^k. I am not sure of the answer, and I leave this as a possible project for the interested reader.)

Later on, we will also need to study infinite collections of balls. The Vitali covering lemma for infinite sets of balls requires a small wrinkle. Suppose that B_i is the ball centered at 0 with radius i. Any two of these balls intersect. So if we take a subset of disjoint balls, it contains only one ball, say B_i. But $3B_i$, or even $100B_i$ does not contain the union of all the balls - which is all of $\mathbb{R}^n$. But the Vitali

covering lemma still holds for infinite collections of balls if we put an extra, slightly technical, hypothesis.

LEMMA 15.14. (Vitali Covering Lemma for infinite collections of balls) Suppose $\{B_i\}_{i\in I}$ is a collection of balls in $\mathbb{R}^n$. Suppose that there is some finite constant M so that any disjoint subset of the balls $\{B_i\}$ has total volume at most M. Then there exists a subcollection $J \subset I$ such that $\{B_j\}_{j\in J}$ are disjoint but $\bigcup_{i\in I} B_i \subset \bigcup_{j\in J} 4B_j$.

PROOF. Any ball B_i has volume at most M, and so the radii of B_i are uniformly bounded. Let B_{j_1} be a ball with radius at least (3/4) times the supremal radius. Add j_1 to J. Among all balls disjoint from B_{j_1}, let B_{j_2} be a ball with radius at least (3/4) times the supremal radius. Add j_2 to J. We continue in this way. The process may terminate in finitely many steps, or we may get a countable sequence of balls B_{j_k}. If the process goes on forever, then the total volume of the balls B_{j_k} is finite (at most M), and so the radius of B_{j_k} tends to zero.

Suppose that $i \notin J$. We have to show that $B_i \subset \cup_{j\in J} 4B_j$. Since i was not added to J, and since the radius of B_{j_k} tends to zero, there must be a first j_k so that $B_i \cap B_{j_k}$ is non-empty. Since B_i is disjoint from all the previous balls of J, the radius of B_{j_k} must be at least (3/4) times the radius of B_i. But then $B_i \subset 4B_{j_k}$. □

15.3.1. Hardy-Littlewood maximal function. Next we consider the averages of functions on various balls. Denote the average of a function f on a set A by

$$\fint_A f := \frac{1}{\mathrm{Vol}A} \int_A f.$$

The *Hardy-Littlewood maximal function* of f is defined by

$$Mf(x) := \sup_r \fint_{B(x,r)} |f|.$$

If f is continuous, then $Mf(x) \geq |f(x)|$. It can happen that $Mf(x) > |f(x)|$. Hardy and Littlewood wanted to understand how much bigger Mf can be than f. For instance, if f is the characteristic function of the unit ball, χ_{B_1}, then it is straightforward to check that

$$M\chi_{B_1}(x) = 1 \text{ if } x \in B_1,$$

$$M\chi_{B_1}(x) \sim |x|^{-n} \text{ if } x \notin B_1.$$

In particular, $\int_{\mathbb{R}^n} M\chi_{B_1}$ is infinite, but $M\chi_{B_1}$ is in L^p for every $p > 1$. Hardy and Littlewood were able to prove the following general estimate about $\|Mf\|_{L^p}$.

THEOREM 15.15. (Hardy and Littlewood) For any dimension n and any $p > 1$, there is a constant $C(n,p)$ so that

$$\|Mf\|_{L^p(\mathbb{R}^n)} \leq C(n,p)\|f\|_{L^p(\mathbb{R}^n)}.$$

This theorem plays an important role in harmonic analysis. The combinatorics of how balls overlap plays a key role in the proof.

Recall that for a function u and a number h,

$$S_u(h) := \{x \in \mathbb{R}^n : |u| > h\}.$$

As in the proof of the Sobolev inequality in Section 15.1, we begin by studying $S_{Mf}(h)$. We study $S_{Mf}(h)$ using the Vitali covering lemma.

LEMMA 15.16. For each $h > 0$, $|S_{Mf}(h)| \lesssim h^{-1}\|f\|_1$.

PROOF. For each $x \in S_{Mf}(h)$, there exists $r(x)$ such that $\fint_{B(x,r(x))} |f| \geq h$, so $\int_{B(x,r(x))} |f| \geq h|B(x,r(x))|$. These $B(x,r(x))$ cover $S_{Mf}(h)$. We wish to apply the Vitali covering lemma to this collection of balls. To apply Lemma 15.14, we need to check that for any disjoint subcollection of these balls, the total volume is at most some number M. But if $\{B_\alpha\}$ is a disjoint subcollection of these balls, then

$$\int_{\mathbb{R}^n} |f| \geq \int_{\cup_\alpha B_\alpha} |f| = \sum_\alpha \int_{B_\alpha} |f| \geq h \sum_\alpha |B_\alpha|.$$

Therefore, $\sum_\alpha |B_\alpha| \leq h^{-1}\|f\|_{L^1}$. With this technical hypothesis checked, we can apply Lemma 15.14. The Lemma guarantees that we can find disjoint B_j's so that $S_{Mf}(h) \subset \cup_j 4B_j$. Hence,

$$|S_{Mf}(h)| \lesssim \sum_j |B_j| \leq h^{-1} \int_{\bigcup B_j} |f| \leq h^{-1}\|f\|_1.$$

□

Our next goal is to use this estimate for $|S_{Mf}(h)|$ to prove Theorem 15.15. One approach would be to consider dyadic h and add their contributions. As in the last section, let $T_{Mf}(k) := \{x \in \mathbb{R}^n : 2^k < |Mf| \leq 2^{k+1}\} \subset S_{Mf}(2^k)$. We note that

$$\int |Mf|^p \sim \sum_{k=-\infty}^{\infty} |T_{Mf}(k)| 2^{kp} \leq \sum_k |S_{Mf}(2^k)| 2^{kp}.$$

If we plug in Lemma 15.16 we get a divergent sum

$$\int |Mf|^p \lesssim \sum_k 2^{-k} 2^{kp} \|f\|_{L^1}.$$

As in the proof of the Sobolev inequality, we need a slight modification of the previous lemma. We observe that if $|f| \leq h/2$ everywhere, then $S_{Mf}(h)$ would be empty. More generally, $|S_{Mf}(h)|$ can be controlled by the integral of $|f|$ over the region where $|f| > h/2$. More precisely:

LEMMA 15.17. $|S_{Mf}(h)| \lesssim h^{-1} \int_{S_f(h/2)} |f|$.

PROOF. In the previous proof, we found disjoint balls B_j so that $4B_j$ covers $S_{Mf}(h)$ and so that for each j,

$$\int_{B_j} |f| \geq h|B_j|.$$

We note that $\int_{B_j \setminus S_f(h/2)} |f| \leq \frac{h}{2}|B_j|$, and so

$$\int_{B_j \cap S_f(h/2)} |f| \geq \frac{h}{2}|B_j|.$$

Therefore, we get

$$|S_{Mf}(h)| \leq |\cup_j 4B_j| \lesssim \sum_j |B_j| \lesssim h^{-1} \sum_j \int_{B_j \cap S_f(h/2)} |f| \leq$$

$$\leq h^{-1} \int_{S_f(h/2)} |f|.$$

□

With this refined Lemma in hand, we can now prove Theorem 15.15

Proof.

$$\int |Mf|^p \lesssim \sum_{k=-\infty}^{\infty} |S_{Mf}(2^k)| 2^{kp}.$$

By Lemma 15.17, the right hand side is

$$\lesssim \sum_k 2^{k(p-1)} \int_{S_f(2^{k-1})} |f|.$$

By interchanging summation and integral, we have

$$= \int_{\mathbb{R}^n} |f| \left(\sum_{2^{k-1} \leq |f|} 2^{k(p-1)} \right). \tag{15.3}$$

Since $p > 1$, the sum is a geometric sum, and the largest term dominates. The largest term occurs when $2^k \sim |f|$, and it has size $\sim |f|^{p-1}$. Therefore, Equation 15.3 is

$$\sim \int |f| \cdot |f|^{p-1} = \int_{\mathbb{R}^n} |f|^p.$$

So, $\|Mf\|_p \lesssim \|f\|_p$.

□

15.3.2. Proof of the Hardy-Littlewood-Sobolev inequality. In this subsection, we give the proof of Theorem 15.10, the Hardy-Littlewood-Sobolev inequality. This inequality gives L^p estimates for the operator T_α. The proof involves some computations and some complicated exponents. (Even the statement of the theorem involves a fairly complicated formula involving the exponents.) We try to emphasize the key ideas in the proof, and we leave some of the computations as exercises for the reader.

Recall that $T_\alpha f(x)$ is defined by

$$T_\alpha f(x) = \int_{\mathbb{R}^n} f(y)|x-y|^{-\alpha} dy.$$

We first observe that $T_\alpha f(x)$ can be computed in terms of the average value of f on $B(x,r)$ for all $0 < r < \infty$.

Lemma 15.18.

$$T_\alpha f(x) = \int_0^\infty r^{n-\alpha-1} \left(\fint_{B(x,r)} f \right) dr.$$

(The proof is a computation which we leave to the reader.)

Next we need some upper bounds for $\oint_{B(x,r)} f$. One upper bound comes from the Hardy-Littlewood maximal function. By the definition of $Mf(x)$ we get immediately

$$\left| \oint_{B(x,r)} f \right| \leq Mf(x). \tag{15.4}$$

We can also get an upper bound in terms of $\|f\|_{L^p}$ by using Holder's inequality.

$$\left| \oint_{B(x,r)} f \right| \lesssim r^{-n} \int_{B(x,r)} |f| \lesssim r^{-n} \|f\|_p r^{n(p-1)/p} = r^{-n/p} \|f\|_p \tag{15.5}$$

Using just one of these two inequalities does not give any finite bound on $\|T_\alpha f\|_{L^q}$. The key idea is to combine these two inequalities:

$$\left| \oint_{B(x,r)} f \right| \lesssim \min \left(Mf(x), r^{-n/p} \|f\|_p \right). \tag{15.6}$$

Plugging this bound into the expression for $T_\alpha f$ in Lemma 15.18, we get

$$|T_\alpha f(x)| \lesssim \int_0^\infty r^{n-\alpha-1} \min \left(Mf(x), r^{-n/p} \|f\|_p \right) dr.$$

It is an exercise to evaluate the right-hand side. When we do so, we get a bound of the form

$$|T_\alpha f(x)| \lesssim |Mf(x)|^A \|f\|_p^B,$$

where A and B depend on α, p, and n.

Exercise 15.5. Do this computation and find $A(\alpha, p, n)$ and $B(\alpha, p, n)$.

Exercise 15.6. Give a conceptual explanation why $A + B = 1$.

We want to bound $\|T_\alpha f\|_{L^q}$ for some q. For any q, we now have the bound

$$\int |T_\alpha f(x)|^q \lesssim \|f\|_p^{Bq} \int |Mf(x)|^{Aq}.$$

We now choose q so that $Aq = p$. With this condition on q, we see that

$$\|T_\alpha f\|_{L^q}^q \lesssim \|f\|_p^{Bq} \|Mf\|_{L^p}^{Aq}.$$

Applying Theorem 15.15, as long as $p > 1$, we get

$$\|T_\alpha f\|_{L^q}^q \lesssim \|f\|_p^{Bq} \|Mf\|_{L^p}^{Aq} \lesssim \|f\|_{L^p}^{(A+B)q} = \|f\|_{L^p}^q.$$

Therefore, we get the bound $\|T_\alpha f\|_{L^q} \lesssim \|f\|_{L^p}$ as long as $p > 1$ and $A(\alpha, p, n)q = p$. After solving for $A(\alpha, p, n)$, the reader can check that this is the condition on the exponents in the statement of the theorem.

15.4. Intersection patterns of tubes in Euclidean space

In this section, we study the possible intersection patterns of cylindrical tubes in Euclidean space. The questions we ask are similar to the ones for balls, but the answers are much more difficult and many of the problems are still open.

We began Section 15.3 by discussing the ball doubling problem. We now pose an analogous problem for tubes: the tube doubling problem. Suppose that $T_i \subset \mathbb{R}^n$ are cylindrical tubes of radius 1 and length N. Let $2T_i$ be the concentric tube of radius 2 and length $2N$ formed by dilating T_i around its center by a factor of 2. For a single tube T, we note that $|2T| = 2^n|T|$. Is there a constant C_n so that for any N, for any set of $1 \times N$ tubes T_i in $\mathbb{R}^n$,

$$|\cup_i 2T_i| \leq C_n |\cup_i T_i|?$$

The answer to this question turns out to be no. Besicovitch gave a beautiful counterexample, in which

$$|\cup_i 2T_i| \sim \frac{\log N}{\log\log N} |\cup_i T_i|.$$

(A refinement by Schoenberg [**Sch**] removes the $\log\log N$ factor.)

For any given value of N, it is not hard to prove an estimate of the form $|\cup_i 2T_i| \leq C_n(N)|\cup_i T_i|$, with a constant $C_n(N)$ depending on N. But the dependence of $C_n(N)$ on N is poorly understood. No one has found a more extreme example than the one constructed by Besicovitch and Schoenberg, leading to the following conjecture:

Conjecture 15.19. (Tube doubling conjecture) For any dimension n, for any $\varepsilon > 0$, there is a constant $C_n(\varepsilon)$, so that the following estimate holds for any N. If T_i are tubes of radius 1 and length N, then

$$|\cup_i 2T_i| \leq C_n(\varepsilon) N^{\varepsilon} |\cup_i T_i|.$$

This conjecture is known in dimension 2, but it is open for all $n \geq 3$.

There are a number of conjectures about tubes in this spirit. The most famous is the Kakeya conjecture. We now formulate one version of the Kakeya conjecture. For a tube $T \subset \mathbb{R}^n$ as above, we write $v(T) \in S^{n-1}$ for a unit vector parallel to the axis of symmetry of T. (There are two choices of $v(T)$, differing by a sign.) We call $v(T)$ the direction of the tube T.

Definition 15.20. Suppose that $T_i \subset \mathbb{R}^n$ are tubes of length N and radius 1. $\{T_i\}$ is a Kakeya set of tubes if $\{v(T_i)\}$ is $\frac{1}{N}$-separated and $\frac{2}{N}$-dense in S^{n-1}.

Our question is: how small can $|\cup T_i|$ be? Each tube T_i has volume $\sim N$, and the number of tubes in a Kakeya set of tubes is $\sim N^{n-1}$. If the tubes were disjoint, then we would have $|\cup_i T_i| \sim N^n$. The construction of Besicovitch mentioned above gives a Kakeya set of tubes in the plane with

$$|\cup_i T_i| \sim \frac{\log\log N}{\log N} N^2.$$

A refinement by Schoenberg [**Sch**] removes the $\log\log N$ factor. In this example, a typical point of $\cup_i T_i$ lies in $\sim \log N$ tubes, giving a compression by a factor $\log N$. These constructions generalize to higher dimensions, giving examples in $\mathbb{R}^n$ with $|\cup_i T_i| \lesssim (\log N)^{-1} N^n$. No one has found an example where $|\cup_i T_i|$ is significantly smaller, leading to the following conjecture.

CONJECTURE 15.21. (Kakeya Conjecture, tube version) In any dimension $n \geq 2$, for any $\varepsilon > 0$, there is a constant $C_{n,\varepsilon}$ so that for any N the following holds. For any Kakeya set of tubes $T_i \subset \mathbb{R}^n$ of dimensions $1 \times N$,

$$|\cup T_i| \geqslant C_{n,\epsilon} \cdot N^{n-\epsilon}.$$

(There are several closely related but not exactly equivalent versions of the Kakeya conjecture. For completeness, we mention the most standard and famous version. Suppose that $K \subset \mathbb{R}^n$ contains a unit line segment in every direction. Then the Hausdorff dimension of K is equal to n.)

Now we give the example of Besicovitch. We will construct a set of N rectangles in the plane, R_j, with width $1/N$ and length 1, with slopes changing evenly between 0 and 1, and with a lot of overlap.

For integers $1 \leq j \leq N$, let $l_j : [0,1] \to \mathbb{R}$ be a list of affine linear functions of the form

$$l_j(x) = \frac{j}{N}x + H(j).$$

Let R_j be the $1/N$ neighborhood of the graph of l_j, which contains a rectangle of width $1/N$ and length 1.

THEOREM 15.22. Suppose that N is an integer of the form A^A for some large integer A. Let R_j be the rectangles described above. If we choose the constants $H(j)$ correctly, then

$$|\cup_j R_j| \lesssim A^{-1} \lesssim \frac{\log\log N}{\log N}.$$

Note that $\sum_j |R_j| \sim 1$, and so this inequality represents a compression by a factor $\frac{\log\log N}{\log N}$.

PROOF. This is a multiscale argument. In order to talk about the different scales, we expand j/N in base A:

$$\frac{j}{N} = \sum_{a=1}^{A} j(a)A^{-a}.$$

In this equation $j(a)$ are the digits in the base A decimal expansion of j/N. The first term $j(1)$, is the first digit after the decimal, and it contributes the largest amount to j/N. The different values of a represent different scales in the problem.

We will choose $H(j)$ so that the following key estimate holds.

PROPERTY 15.23. Suppose $1 \leq b \leq A$. If $j(a) = j'(a)$ for $1 \leq a \leq b-1$, then for all $x \in \left[\frac{A-b}{A}, \frac{A-b+1}{A}\right]$,

$$|l_j(x) - l_{j'}(x)| \leq 4A^{-b}.$$

Property 15.23 quickly implies our desired bound on $|\cup_j R_j|$. Fix a value of b. For a given choice of $j(1), ..., j(b-1)$, consider all the rectangles $R_{j'}$ where $j'(a) = j(a)$ for $1 \leq a \leq b-1$. By Property 15.23, each $R_{j'} \cap \left[\frac{A-b}{A}, \frac{A-b+1}{A}\right] \times \mathbb{R}$ lies in the parallelogram defined by the inequalities

$$\frac{A-b}{A} \leq x \leq \frac{A-b+1}{A} \text{ and } |y - l_j(x)| \leq 5A^{-b}.$$

This parallelogram has area $10A^{-b-1}$. Since there are A^{b-1} possible values of $j(1), ..., j(b-1)$, we see that

$$\left| (\cup_j R_j) \cap \left(\left[\frac{A-b}{A}, \frac{A-b+1}{A} \right] \times \mathbb{R} \right) \right| \leq 10A^{-2}.$$

Summing over all choices of b in the range $1 \leq b \leq A$, we see that

$$|(\cup_j R_j) \cap ([0,1] \times \mathbb{R})| \leq 10A^{-1}.$$

The rectangles R_j may stick out a little from $[0,1] \times \mathbb{R}$, but it's straightforward to see that $|\cup_j R_j| \leq 11A^{-1}$. So it only remains to choose $H(j)$ in order to guarantee Property 15.23.

We will write $H(j)$ as a sum of A different terms with different orders of magnitude. Each term is designed to arrange Property 15.23 for a particular value of b. We write $H(j)$ as

$$H(j) = \sum_{a=1}^{A} h(a) j(a) A^{-a},$$

where $h(a) \in [-1,1]$ is a constant that we can choose later. The constant $h(a)$ will be designed to make Property 15.23 hold for $b = a$.

Plugging this expression into the formula for $l_j(x)$, we see that

$$l_j \left(\frac{A-b}{A} \right) = \frac{A-b}{A} \cdot \frac{j}{N} + H(j) = \sum_{a=1}^{A} \left(\frac{A-b}{A} + h(a) \right) j(a) A^{-a}.$$

Therefore,

$$\left| l_j \left(\frac{A-b}{A} \right) - l_{j'} \left(\frac{A-b}{A} \right) \right| \leq \sum_{a=1}^{A} \left| \frac{A-b}{A} + h(a) \right| |j(a) - j'(a)| \, A^{-a}.$$

For the rest of the proof, let us suppose that $j(a) = j'(a)$ for $1 \leq j \leq b-1$. These equalities imply that, in the last inequality, the first $b-1$ terms on the right-hand side vanish. We now choose $h(a) := -\frac{A-a}{A}$. With this choice, the b^{th} term on the right-hand side vanishes also. We can easily bound the other terms by noting that $|h(a)| \leq 1$ and $|j(a) - j'(a)| \leq A$, yielding

$$\left| l_j \left(\frac{A-b}{A} \right) - l_{j'} \left(\frac{A-b}{A} \right) \right| \leq \sum_{a=b+1}^{A} A^{-a+1} \leq 2A^{-b}.$$

Now for any $x \in [\frac{A-b}{A}, \frac{A-b+1}{A}]$, we see that

$$|l_j(x) - l_{j'}(x)| \leq \left| l_j \left(\frac{A-b}{A} \right) - l_{j'} \left(\frac{A-b}{A} \right) \right| + \left| \frac{j}{N} - \frac{j'}{N} \right| \left| x - \frac{A-b}{A} \right| \leq$$
$$\leq 2A^{-b} + A^{-b+1} A^{-1} = 3A^{-b}.$$

This finishes the proof of Property 15.23 and hence the proof of Theorem 15.22. □

In this theorem, the directions $v(R_j) \in S^1$ are pairwise separated by $\gtrsim 1/N$. They form a $\sim 1/N$ net for an eighth of the circle S^1. Taking eight rotated copies of this set of rectangles, and rescaling by a factor of N, we get a Kakeya set of $1 \times N$ tubes with total area $\lesssim \frac{\log \log N}{\log N} N^2$.

This example can also be used as a counterexample for the tube doubling problem that we discussed at the beginning of the section. For a rectangle R_j as

above let $v(R_j)$ be the direction of R_j. Recall that $v(R_j)$ is only defined up to sign, and we make the choice so that the x-component of $v(R_j)$ is negative. Now we define R_j^+ to be the translation of R_j by $10v(R_j)$. We note that since R_j has length ~ 1, R_j^+ is contained in $100R_j$.

We claim that for the values of $H(j)$ constructed in the proof above, the rectangles R_j^+ are disjoint. We state this refined result as a corollary.

COROLLARY 15.24. Suppose that N is an integer of the form A^A for some large integer A. Let R_j be the rectangles described in Theorem 15.22. If we choose the constants $H(j)$ correctly, then

$$|\cup_j R_j| \lesssim \frac{\log\log N}{\log N} \sum_j |R_j|,$$

and yet R_j^+ are disjoint.

PROOF. Let $H(j)$ be as in the proof of Theorem 15.24. Recall that we expanded j/N in base A as

$$\frac{j}{N} = \sum_{a=1}^{A} j(a) A^{-a}.$$

In terms of the digits $j(a)$, we define

$$H(j) := -\sum_{a=1}^{A} \frac{A-a}{A} j(a) A^{-a}.$$

A simple calculation shows that for any $1 \le j < j' \le N$, $H(j') < H(j)$.

Therefore, for $x \le -5$, and $1 \le j < j' \le N$, we have

$$l_{j'}(x) - l_j(x) = \left(\frac{j'}{N} - \frac{j}{N}\right) x + (H(j') - H(j)) \le -5 \left|\frac{j'}{N} - \frac{j}{N}\right| \le -5/N.$$

Therefore, R_j^+ and $R_{j'}^+$ are disjoint.

□

From this corollary we see that

$$|\cup_j 100R_j| \ge |\cup_j R_j^+| = \sum_j |R_j| \gtrsim \frac{\log N}{\log\log N} |\cup_j R_j|.$$

With a little bit more care, it is also possible to choose $1 \times N$ rectangles R_j so that $|\cup_j 2R_j| \gtrsim \frac{\log N}{\log\log N} |\cup_j R_j|$.

Besicovitch was interested in these problems about overlapping tubes as natural variations of the problems about overlapping balls that we discussed in the previous section. For example, the Vitali covering lemma is used to prove the Lebesgue differentiation theorem (see for example Chapter 3 of [**StSh**]). Besicovitch asked whether there was a generalization of the Lebesgue differentiation theorem using tubes instead of balls, and he gave a counterexample to such a generalization using the construction above.

In the 1970's, Fefferman found a connection between the Besicovitch construction and oscillatory integral operators. We will discuss this connection in the next section. This discovery led to renewed interest in quantitative estimates for the Kakeya problem and the tube doubling problem.

15.5. Oscillatory integrals and the Kakeya problem

In this section, we study an oscillatory integral operator, and we see how the L^p estimates for such an operator connect with intersection patterns of tubes in Euclidean space.

Earlier in the chapter, we proved the Hardy-Littlewood-Sobolev theorem, Theorem 15.10. This theorem gave L^p estimates for the operator T_α defined by

$$T_\alpha f = f * K_\alpha,$$

where

$$K_\alpha(x) := |x|^{-\alpha}.$$

Note that the kernel K_α is positive. Now we consider a variation of T_α using an oscillating kernel which is sometimes positive and sometimes negative.

$$\tilde{T}_\alpha f := f * \tilde{K}_\alpha,$$

where

$$\tilde{K}_\alpha(x) := [1 + |x|]^{-\alpha} \cos |x|.$$

The function $\tilde{K}_\alpha(x)$ is still radial. Near the origin, it is bounded instead of going to infinity. For $|x| \geq 1$, $|\tilde{K}_\alpha(x)| \lesssim K_\alpha(x)$, and they are often comparable. But $\tilde{K}_\alpha(x)$ oscillates with the radius, so that it has positive and negative parts. If one dropped a stone into a pond and looked at the ripples, the shape would be a little bit like $\tilde{K}_\alpha$, with a modest peak in the center, and then waves going outward and getting smaller the farther they are from the center.

As in the Hardy-Littlewood-Sobolev inequality, we will focus on the range $0 < \alpha < n$. Our main question is: what are all the L^p estimates obeyed by $\tilde{T}_\alpha$?

At first sight, this problem may look like a small variation on the Hardy-Littlewood-Sobolev problem - it is just a similar kernel with some oscillations added. For a reader who has not studied this area before, it may be surprising to learn that this is a major open problem of analysis. The kernel $\tilde{K}_\alpha$ has positive and negative parts, and so in the convolution $\tilde{T}_\alpha f = f * \tilde{K}_\alpha$, some cancellation can occur. The key issue is to understand how much cancellation needs to occur. This problem turns out to be connected to the intersection patterns of long thin tubes in $\mathbb{R}^n$. This may be a little surprising, since the kernel $\tilde{K}_\alpha$ is not shaped like a long thin tube. The goal of this section is to explain this connection.

We will focus on estimates of the form $\|\tilde{T}_\alpha f\|_p \lesssim \|f\|_p$, so that we have fewer parameters to keep track of. (Estimates of the form $\|\tilde{T}_\alpha f\|_q \lesssim \|f\|_p$ are interesting also, but all of the essential issues already appear in the case $q = p$.)

Let us consider how the operator $\tilde{T}_\alpha$ behaves on some examples. When we studied the Hardy-Littlewood-Sobolev inequality, the key examples were characteristic functions of balls. So let us consider the case $f = \chi_{B_r}$ for various radii r. We have

$$\tilde{T}_\alpha \chi_{B_r}(x) = \int_{B_r} \tilde{K}_\alpha(x - y) dy = \int_{B_r} [1 + |x - y|]^{-\alpha} \cos |x - y| dy. \tag{15.7}$$

Even in this case, it's not trivial to estimate $|\tilde{T}_\alpha \chi_{B_r}(x)|$, because of the cancellation in the integral. The easiest case to understand is the case of small r, say $r = 1/100$. In this case, we see that at most points x, the sign of $\cos |x - y|$ is

constant for $y \in B_r$, and $|\cos(x-y)| \sim 1$ for $y \in B_r$. At such a point x, we get

$$|\tilde{T}_\alpha \chi_{B_r}(x)| \sim \int_{B_r} [1+|x-y|]^{-\alpha}\, dy \sim |\tilde{K}_\alpha(x)|.$$

If we define $f_1 := \chi_{B_{1/100}}$, then for most x, $|\tilde{T}_\alpha f_1(x)| \sim |\tilde{K}_\alpha(x)|$. For any p, we have $\|f_1\|_p \sim 1$. On the other hand, $\int |\tilde{T}_\alpha f_1|^p \sim \int_{\mathbb{R}^n} (1+|x|)^{-\alpha p}$ is finite if and only if $\alpha p > n$.

Considering $r < 1/100$ doesn't give any new information. Now we turn to larger scales $r \gg 1$. For a function supported on a ball of large radius r, there is a cleverer choice than χ_{B_r}. Suppose that we want to make $\tilde{T}_\alpha f(0)$ large. Let us write it out as an integral:

$$\tilde{T}_\alpha f(0) = \int_{\mathbb{R}^n} f(y)[1+|y|]^{-\alpha} \cos|y| dy.$$

If we choose f carefully, then all the contributions in the integral are positive, instead of cancelling each other. This motivates defining

$$f_2 := \chi_{B_r} \operatorname{Sign}(\cos|y|).$$

Here f_2 also depends on r, and it will be interesting for large $r \to \infty$. We have $\|f_2\|_p = r^{n/p}$. We also have

$$|\tilde{T}_\alpha f_2(0)| = \int_{B_r} [1+|y|]^{-\alpha} |\cos|y||dy \sim r^{n-\alpha}.$$

In fact, for all $|x| < 1/100$, we have $|\tilde{T}_\alpha f_2(x)| \sim r^{n-\alpha}$. This requires a little more thought, and we state it as an exercise.

EXERCISE 15.7. Suppose that $r \geq 1$. For all $|x| < 1/100$, prove that $|\tilde{T}_\alpha f_2(x)| \sim r^{n-\alpha}$. Here is some intuition why this bound is true. If we write out the definition of $\tilde{T}_\alpha$ and the definition of f_2, we get

$$\tilde{T}_\alpha f_2(x) = \int f_2(y) \tilde{K}_\alpha(x-y) dy = \int_{B_r} \operatorname{Sign}(\cos|y|)[1+|x-y|]^{-\alpha} \cos|x-y| dy.$$

We reorganize this last expression so that all the possibly negative terms are at the end:

$$= \int_{B_r} [1+|x-y|]^{-\alpha} |\cos|x-y|| \operatorname{Sign}(\cos|y|) \operatorname{Sign}(\cos|y-x|) dy.$$

The integrand is positive as long as $\operatorname{Sign}(\cos|y|) = \operatorname{Sign}(\cos|y-x|)$. Since $x < 1/100$ is small, this equality holds for almost all y. Therefore, morally, our integral should be similar to

$$\int_{B_r} [1+|x-y|]^{-\alpha} |\cos|x-y||dy \sim \int_{B_r} [1+|x-y|]^{-\alpha} dy \sim r^{n-\alpha}.$$

The exercise is to make this intuition rigorous.

Using the last exercise, we see that $\|\tilde{T}_\alpha f_2\|_p \gtrsim r^{n-\alpha}$. We saw above that $\|f_2\|_p \sim r^{n/p}$. Therefore $\|\tilde{T}_\alpha f_2\|_p \lesssim \|f_2\|_p$ if and only if $n/p \geq n-\alpha$.

Combining the information we got from the example f_1 and the example f_2, we have the following proposition.

PROPOSITION 15.25. If $\|\tilde{T}_\alpha f\|_p \lesssim \|f\|_p$ for the examples above, f_1 and f_2, then

$$\frac{n}{\alpha} < p \leq \frac{n}{n-\alpha}.$$

EXERCISE 15.8. There is a slightly stronger version of the example f_2. Define $f_3 = \chi_{B_r} \tilde{K}_{n-\alpha}$. Estimate $\|\tilde{T}_\alpha f_3\|_p$ and $\|f_3\|_p$, and check that if $\|\tilde{T}_\alpha f_2\|_p \lesssim \|f_2\|_p$ (for all r), then $n/p > n - \alpha$. Therefore, the estimate on p in Proposition 15.25 can be improved to $\frac{n}{\alpha} < p < \frac{n}{n-\alpha}$.

The kernel $\tilde{K}_\alpha$ is spherically symmetric. So far we have been considering $\tilde{T}_\alpha f = f * \tilde{K}_\alpha$ for spherically symmetric functions f. This might seem natural, but it turns out that functions f that are far from spherically symmetric play an important role.

The next function we consider is an oscillating function supported on a long thin tube. This function plays a crucial role in the theory, and it explains why long thin tubes should be relevant to studying $\tilde{T}_\alpha$.

Let T be a cylinder of length $L >> 1$ and radius $(1/1000)L^{1/2}$. The cylinder may point in any direction. Let v_T be a unit vector parallel to the axis of the cylinder. Let f_T be the function

$$f_T(x) := \chi_T(x) e^{i(v_T \cdot x)}.$$

Let T^+ denote the cylinder we get by translating T by $10Lv_T$. We are going to study the behavior of $\tilde{T}_\alpha f_T$ on T^+.

PROPOSITION 15.26. Fix a dimension $n \geq 2$. For all L sufficiently large, the following holds. If f_T and T^+ are defined as above, then for every $x \in T^+$ we have

$$|\tilde{T}_\alpha f_T(x)| \gtrsim L^{\frac{n+1}{2} - \alpha}.$$

As a corollary, we get new information about the possible bounds of the form $\|\tilde{T}_\alpha f\|_p \lesssim \|f\|_p$.

COROLLARY 15.27. If $\alpha < \frac{n+1}{2}$, then for every $p \in [1, \infty]$, as $L \to \infty$,

$$\frac{\|\tilde{T}_\alpha f_T\|_p}{\|f_T\|_p} \to \infty.$$

So the operator $\tilde{T}_\alpha$ does not obey any L^p estimate of the form $\|\tilde{T}_\alpha f\|_p \lesssim \|f\|_p$.

PROOF. Notice that T^+ has the same size as T. The function f_T has size ~ 1 and support on T. If $\alpha < \frac{n+1}{2}$, then the function $\tilde{T}_\alpha f_T$ has size $>> 1$ on T^+. So $\|\tilde{T}_\alpha f_T\|_p \sim L^{\frac{n+1}{2} - \alpha} \|f_T\|_p$. □

Before the proof of Proposition 15.26, we give some intuition. Consider a point x in T^+.

$$\tilde{T}_\alpha f_T(x) = \int_T |x-y|^\alpha \cos|x-y| e^{i(v_t \cdot y)} dy.$$

Now the key point is that the oscillations of $e^{i(v_t \cdot y)}$ and the oscillations of $\cos|x-y|$ are in sync on T. Without any real loss of generality, we can choose coordinates so that $x = 0$. To get some intuition how $e^{i(v_t \cdot y)}$ and $\cos|x-y| = \cos|y|$ interact, we visualize the set where $e^{i(v_t \cdot y)}$ is equal to 1 and the set where $\cos|y| = 1$, and we will see that these two sets are close to each other. We first consider the set where $e^{iv_t \cdot y}$ is equal to 1. The set $\{y | e^{iv_t \cdot y} = 1\} = \{y | v_t \cdot y \in 2\pi\mathbb{Z}\}$ is a union of parallel planes, perpendicular to v_T, with spacing 2π between them. The set

$\{y|\cos|y| = 1\} = \{y||y| \in 2\pi\mathbb{Z}\}$ is a union of concentric spheres around $x = 0$ with spacing 2π between them. We illustrate these planes and spheres in Figure 15.1. As long as T is narrow enough, the spheres and planes nearly coincide inside of T. We will make this idea quantitative in the proof below, using the Pythagorean theorem.

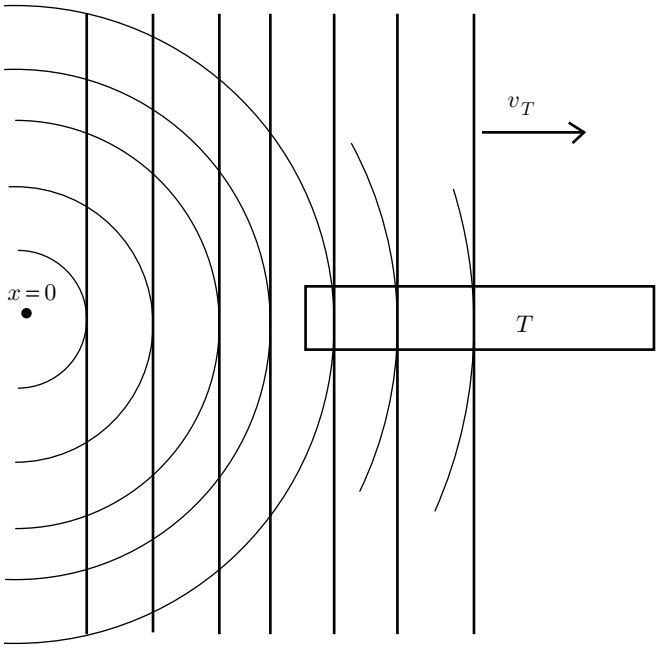

FIGURE 15.1. Oscillations of $\cos|y|$ are in sync with $e^{iv_t \cdot y}$ on T.

PROOF OF PROPOSITION 15.26. Recall that $x \in T^+$ and

$$\tilde{T}_\alpha f_T(x) = \int_T |x-y|^\alpha \cos|x-y| e^{i(v_t \cdot y)} dy. \tag{15.8}$$

We decompose the vector $x-y$ into a component parallel to v_T and a component perpendicular to v_T. Because of the geometry of T and T^+, the v_t component of $x-y$ has length $\geq 5L$, and the perpendicular component has length $\leq (1/1000)L^{1/2}$. By the Pythagorean theorem, we have

$$(v_t \cdot x - v_t \cdot y)^2 \leq |x-y|^2 \leq (v_t \cdot x - v_t \cdot y)^2 + 10^{-6}L.$$

Since $|v_t \cdot x - v_t \cdot y| \geq 5L$, we see that

$$\big||x-y| - |v_t \cdot x - v_t \cdot y|\big| \leq 10^{-6}.$$

Since $|\cos a - \cos b| \leq |a-b|$, and since $\cos b = \cos(-b)$, we see that

$$\big|\cos|x-y| - \cos(v_t \cdot x - v_t \cdot y)\big| \leq 10^{-6}.$$

Plugging this estimate into Equation 15.8, we get

$$\tilde{T}_\alpha f_T(x) = \int_T |x-y|^\alpha \cos(v_t \cdot x - v_t \cdot y) e^{i(v_t \cdot y)} dy + \text{ small error},$$

where

$$|\text{small error}| \leq 10^{-6} \int_T |x-y|^{\alpha} dy.$$

Expanding $\cos a = \frac{1}{2}(e^{ia} + e^{-ia})$, we get

$$\tilde{T}_\alpha f_T(x) = \frac{1}{2} e^{iv_t \cdot x} \int_T |x-y|^{-\alpha} dy + \frac{1}{2} e^{-iv_t \cdot x} \int_T |x-y|^{-\alpha} e^{2iv_t \cdot y} dy + \text{ small error.}$$

The first integral is the main term. We see that it dominates the small error. The second term has a lot of cancellation in it (at least for L large), and so it's not hard to check that the first term also dominates the second term. Therefore, $|\tilde{T}_\alpha f_T(x)| \sim \int_T |x-y|^{-\alpha} dy$. The volume of T is $\sim L^{\frac{n+1}{2}}$, and $|x-y| \sim L$ for $y \in T$, and so

$$|\tilde{T}_\alpha f_T(x)| \sim \int_T |x-y|^{-\alpha} dy \sim L^{\frac{n+1}{2} - \alpha}.$$

□

This type of example, an oscillating function supported on a long thin tube, plays an important role for studying several linear operators in Fourier analysis and partial differential equations. For instance, there are similar examples connected to the wave equation. In this book, we won't discuss the mathematics of the wave equation, but we spend a paragraph trying to describe in words how such examples might occur for sound waves.

Imagine an airplane traveling at the speed of sound. The path of the airplane in space-time is like a long thin tube. The engine of the plane vibrates, making sound waves, and these sound waves travel at the same speed as the airplane. At each moment of time, the airplane experiences the sound waves that the engine generated at every previous moment of time. Because of this effect, the airplane experiences dramatically stronger sound waves than it would have felt at a lower or higher speed. (This accumulation of waves is one of the engineering challenges associated with airplanes reaching the speed of sound.) Even if the airplane turns off the engine and coasts at the speed of sound, it will continue to experience strong sound waves for some time, since the sound waves generated by the engine when it was on will still be moving with the airplane. If we consider the air pressure as a function of space and time, then we will see high amplitude vibrations in the region near the path of the airplane - a long thin tube. There is a linear operator W that has input f, the forces generating sound waves - in this case the engine - and has output Wf, the resulting air pressure. In our case, the action of the engine, modeled by a function f, is supported on a tube in space time, around the path of the airplane, stopping at a certain time t_0 when the engine is turned off. The air pressure, Wf, experiences strong vibrations on a longer tube, extending after the engine has been turned off. This situation is analogous to Proposition 15.26. So although the operator $\tilde{T}_\alpha$ is not an accurate model for sound waves, the mathematical issues in understanding it are similar to those in the wave equation. See [**Ta1**] and the references there for more introduction to these issues.

We now return to our operators $\tilde{T}_\alpha$. Because of the tube example, we saw that for $\alpha < \frac{n+1}{2}$, there are no L^p estimates for $\tilde{T}_\alpha$. For $\alpha \geq \frac{n+1}{2}$, all the examples we have seen so far obey the inequality $\|\tilde{T}_\alpha f\|_p \lesssim \|f\|_p$ for all p in the range $\frac{n}{\alpha} < p < \frac{n}{n-\alpha}$.

We now focus on the case $\alpha = \frac{n+1}{2}$. Fourier anaysts worked hard on this case from the 1930's through the 1960's. Until the early 70's, it was generally believed that $\|\tilde{T}_{\frac{n+1}{2}} f\|_p \lesssim \|f\|_p$ for all p in the range $(\frac{n}{\alpha}, \frac{n}{n-\alpha}) = (\frac{2n}{n+1}, \frac{2n}{n-1})$, although mathematicians could only prove the inequality in the very special case $p = 2$. In the early 70's, Fefferman gave a counterexample, showing that the inequality $\|\tilde{T}_{\frac{n+1}{2}} f\|_p \lesssim \|f\|_p$ is false for all $p \neq 2$ ([**Fef**]). This counterexample is one of the most surprising and interesting in the theory of linear operators.

THEOREM 15.28. ([**Fef**]) $\|\tilde{T}_{\frac{n+1}{2}} f\|_p \lesssim \|f\|_p$ is false for all $p \neq 2$.

PROOF. We will prove the theorem for $p > 2$, and then indicate how to modify the argument to deal with $p < 2$.

The main idea of the construction is to let f be a sum of many functions f_{T_i} where the T_i are long thin tubes arranged according to Besicovitch's construction. The geometry of the interesecting tubes will influence the L^p norms of f and $\tilde{T}_{\frac{n+1}{2}} f$.

We define $f = \sum_i f_{T_i}$ where T_i are tubes of length L and radius $\frac{1}{1000} L^{1/2}$, and $v(T_i)$ is a unit vector pointing in the direction of T_i. As above, we define each f_T by:

$$f_T(x) := \chi_T(x) e^{i(v_T \cdot x)}.$$

Since $\tilde{T}_\alpha$ is linear, we have

$$\tilde{T}_\alpha f = \sum_i \tilde{T}_\alpha f_{T_i}.$$

Using Besicovitch's construction, we can arrange that the tubes T_i are disjoint and yet the tubes T_i^+ intersect heavily. We essentially proved this in two dimensions in Corollary 15.24 above. The result can easily be generalized to higher dimensions.

THEOREM 15.29. (Besicovitch, 1920's) Fix a dimension $n \geq 2$. For any $L \geq 1$, there is a finite set of disjoint tubes T_i (with length L and radius $\sim (1/1000) L^{1/2}$), with the property that

$$|\cup_i T_i^+| = \mu(L)^{-1} |\cup_i T_i|,$$

where $\mu(L) \gtrsim \frac{\log L}{\log \log L} \to \infty$.

PROOF SKETCH. Starting with Corollary 15.24, switch the roles of R_j and R_j^+ and rescale the rectangles so that they have dimensions $(1/1000) L^{1/2} \times L$. In higher dimensions, just thicken each two-dimensional rectangle R_j to an n-dimensional tube T_j. □

We define $f = \sum_i f_{T_i}$ where $\{T_i\}$ is the set of tubes coming from this theorem, illustrated in Figure 15.2. We now want to estimate $\|f\|_p$ and $\|\tilde{T}_{\frac{n+1}{2}} f\|_p$. We can easily determine $\|f\|_p$. Recall that $|f_{T_i}(x)| = \chi_{T_i}(x)$. Since the tubes T_i are disjoint, we get

$$\|f\|_{L^p}^p = \sum_i |T_i|.$$

Next we want to prove a lower bound for $\|\tilde{T}_{\frac{n+1}{2}} f\|_p$. We first give a heuristic argument, and then make it rigorous. By Proposition 15.26, we know that for all $x \in T_i^+$,

$$|\tilde{T}_{\frac{n+1}{2}} f_{T_i}(x)| \gtrsim 1.$$

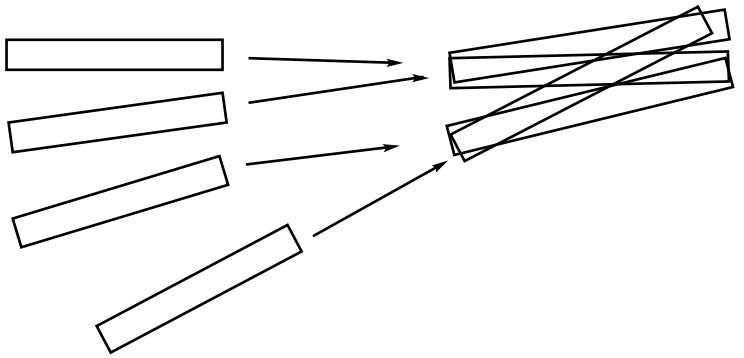

FIGURE 15.2. Tubes T_i (left) are disjoint. Tubes T_i^+ (right) overlap heavily in Besicovitch construction. Arrows point from T_i to T_i^+.

A typical point in $X = \cup_i T_i^+$ lies in $\gtrsim \mu(L)$ different tubes T_i^+. Therefore, $\sum_i \tilde{T}_{\frac{n+1}{2}} f_{T_i}(x)$ includes $\gtrsim \mu(L)$ summands of norm $\gtrsim 1$. Recall that the summands are complex numbers that can point in any direction. How big do we expect the sum to be? There could well be cancellation, so it's not actually reasonable to expect $|\sum_i \tilde{T}_{\frac{n+1}{2}} f_{T_i}(x)| \sim \sum_i |\tilde{T}_{\frac{n+1}{2}} f_{T_i}(x)|$. An important reference point is the sum of random numbers. If a_i is a finite list of complex numbers, and we consider the sum $\sum_i \pm a_i$ with random signs, then with high probability $|\sum_i \pm a_i| \sim (\sum_i |a_i|^2)^{1/2}$. We will state a more precise version of this inequality in a moment. Unless we especially craft f_{T_i} to avoid it, it is reasonable to expect this type of square root cancellation. So at a typical point $x \in \cup_i T_i^+$, we expect

$$|\sum_i \tilde{T}_{\frac{n+1}{2}} f_{T_i}(x)| \sim \left(\sum_i |\tilde{T}_{\frac{n+1}{2}} f_{T_i}(x)|^2 \right)^{1/2} \gtrsim \mu(L)^{1/2}.$$

Therefore, as a heuristic, we expect

$$\|\tilde{T}_{\frac{n+1}{2}} f\|_p^p = \int \left(\sum_i \tilde{T}_{\frac{n+1}{2}} f_{T_i} \right)^p \gtrsim \mu(L)^{p/2} |\cup_i T_i^+| \sim \mu(L)^{\frac{p}{2}-1} \sum_i |T_i|.$$

Since, $\|f\|_{L^p}^p = \sum_i |T_i|$, we get $\|\tilde{T}_{\frac{n+1}{2}} f\|_p^p \gtrsim \mu(L)^{\frac{p}{2}-1} \|f\|_{L^p}^p$, and so for all $p > 2$,

$$\frac{\|\tilde{T}_{\frac{n+1}{2}} f\|_p}{\|f\|_p} \to \infty.$$

Now we start to make this argument rigorous. To get rigorous lower bounds, it is convenient to introduce random signs. We define

$$f_{ran} = \sum_i \pm f_{T_i},$$

where the $\pm$ signs are selected at random. Now we consider the average value of $\|f_{ran}\|_p$ and the average value of $\|\tilde{T}_{\frac{n+1}{2}} f_{ran}\|_p$. Regardless of the signs, we have $\|f_{ran}\|_p = \sum_i |T_i|$.

PROPOSITION 15.30. (Khintchin) If g_i are any functions, and $g = \sum_i \pm g_i$ with random signs, then for any $1 \leq p < \infty$, the average value of $\|g\|_p$ is $\sim_p \|(\sum_i |g_i|^2)^{1/2}\|_p$.

(See [**Wo3**] for a proof of the inequality. We only use an easy direction of the inequality which is proven in Exercise 15.9.)

In particular, the average value of $\|\tilde{T}_{\frac{n+1}{2}} f_{ran}\|_p = \|\sum_i \pm \tilde{T}_{\frac{n+1}{2}} f_{T_i}\|_p$ is

$$\sim \left(\int \left(\sum_i |\tilde{T}_{\frac{n+1}{2}} f_{T_i}|^2 \right)^{p/2} \right)^{1/p}.$$

Since $|\tilde{T}_{\frac{n+1}{2}} f_{T_i}| \gtrsim 1$ on T_i^+, we get

$$\int \left(\sum_i |\tilde{T}_{\frac{n+1}{2}} f_{T_i}|^2 \right)^{p/2} \gtrsim \int (\sum_i \chi_{T_i^+})^{p/2}.$$

The average value of $(\sum_i \chi_{T_i^+})$ on $\cup_i T_i^+$ is $\mu(L)$. Using Holder's inequality, we see that

$$\int (\sum_i \chi_{T_i^+})^{p/2} \gtrsim \mu(L)^{p/2} |\cup_i T_i^+| \sim \mu(L)^{\frac{p-2}{2}} \sum_i |T_i^+|.$$

Combining these inequalities, we see that the average value of $\|\tilde{T}_{\frac{n+1}{2}} f_{ran}\|_p$ is

$$\gtrsim \mu(L)^{\frac{p-2}{2p}} \|f_{ran}\|_p.$$

Since $p > 2$, and $\mu(L) \to \infty$, we see that there is no inequality of the form $\|\tilde{T}_{\frac{n+1}{2}} f\|_p \lesssim \|f\|_p$.

□

We make some brief comments on the case $p < 2$. In this case, we choose the tubes T_i so that T_i overlap a lot and T_i^+ are disjoint. We again take $f = f_{ran} = \sum_i \pm f_{T_i}$. In this case, we see immediately that $\|\tilde{T}_{\frac{n+1}{2}} f_{ran}\|_p^p \gtrsim \sum_i |T_i|$. On the other hand, each point in $\cup_i T_i$ typically lies in $\mu(L)$ different tubes T_i. Therefore, we typically expect $|f_{ran}(x)| \sim \mu(L)^{1/2}$ for $x \in \cup_i T_i$, and hence

$$\|f_{ran}\|_p^p \lesssim \mu(L)^{p/2} |\cup_i T_i| \sim \mu(L)^{\frac{p-2}{2}} \sum_i |T_i|.$$

Since $p < 2$, $\mu(L)^{\frac{p-2}{2}} \to 0$, and we see that $\|\tilde{T}_{\frac{n+1}{2}} f_{ran}\|_p$ is (usually) far bigger than $\|f_{ran}\|_p$. As above this argument can be made rigorous using Khintchin's inequality.

EXERCISE 15.9. Khintchin's inequality says that if g_i are any functions, and $g = \sum_i \pm g_i$ with random signs, then for any $1 \leq p < \infty$, the average value of $\|g\|_p$ is $\sim_p \|(\sum_i |g_i|^2)^{1/2}\|_p$. In our proof of Theorem 15.28, we only actually used one direction of this inequality. Assuming $p > 2$, prove that

$$\text{Avg}_{\text{choice of signs}} \int \left| \sum_i \pm g_i \right|^p \geq \int \left(\sum_i |g_i|^2 \right)^{p/2}.$$

Finally we come to the case $\alpha > \frac{n+1}{2}$. In this case, determing all the L^p estimates for the operator $\tilde{T}_\alpha$ is a major open problem of harmonic analysis. Working with slightly different operators, Bochner and Riesz made a conjecture about these bounds in the 1930's.

CONJECTURE 15.31. Suppose that $n > \alpha > \frac{n+1}{2}$ and $\frac{n}{\alpha} < p < \frac{n}{n-\alpha}$. Then the operator $\tilde{T}_\alpha$ obeys the estimate $\|\tilde{T}_\alpha f\|_p \lesssim \|f\|_p$.

For $\alpha > \frac{n+1}{2}$, the tube examples and the Besicovitch construction are not as dangerous. By Proposition 15.26, if $x \in T^+$, then $|\tilde{T}_\alpha f_T(x)| \sim L^{\frac{n+1}{2}-\alpha}$. For comparison, $|f_T| = 1$ on T. So when $\alpha > \frac{n+1}{2}$, we get a damping effect: $|\tilde{T}_\alpha f_T|$ on T^+ is smaller than $|f_T|$ on T by a power of L. The Besicovitch construction gives a compression by a factor $\mu(L)$ on the order of $\log L$, which is not strong enough to overcome this polynomial damping. However, if there were a generalization of the Besicovitch construction with compression factor at least L^γ for some $\gamma > 0$, then using this construction in the setup $f_{ran} = \sum_i \pm f_{T_i}$ would disprove Conjecture 15.31. In other words, Conjecture 15.31 implies the tube doubling conjecture, Conjecture 15.19.

The theme of this chapter is the interplay between analysis, geometry, and combinatorics on Euclidean space. Earlier, we studied the Hardy-Littlewood-Sobolev operator T_α, and we saw that the L^p estimates for T_α are closely related to estimates about the intersection patterns of balls in Euclidean space. Then we turned to the Bochner-Riesz operators, $\tilde{T}_\alpha$, and we saw that the L^p estimates for $\tilde{T}_\alpha$ are closely related to estimates about the intersection patterns of tubes in Euclidean space.

These problems about intersection patterns of tubes in Euclidean space are difficult open problems. In the early 90's, Bourgain made some partial progress on the Kakeya problem and used it to prove new results about oscillatory integral operators like $\tilde{T}_\alpha$. The best known results about $\tilde{T}_\alpha$ depend on our partial progress about the intersection patterns of tubes. We turn to quantitative estimates about tubes in the next section.

15.5.1. Curvature in Fourier analysis. The issues we described in this section play a role in many problems in Fourier analysis. We won't try to discuss all of these problems in this chapter, but there is one other problem I would like to at least mention because it has played such an important role in the development of the field. This problem is the restriction problem raised by Stein in the late 60's cf. [**St**].

A basic problem in Fourier analysis is to determine all of the L^p inequalities that the Fourier transform obeys: to find all the inequalities of the form

$$\|\hat{f}\|_{L^q(\mathbb{R}^n)} \leq C\|f\|_{L^p(\mathbb{R}^n)}. \tag{15.9}$$

This problem was solved in the early 20th century. The Hausdorff-Young inequalities give all of the inequalities of this form – see the first lecture in [**Ta2**] for details.

The restriction problem concerns a generalization of this problem where we replace $L^q(\mathbb{R}^n)$ on the left-hand side with the L^q norm on a surface $S \subset \mathbb{R}^n$, such as the unit sphere. If $S \subset \mathbb{R}^n$ is a submanifold, and $f : S \to \mathbb{C}$ is a function, then

we can define the L^q norm of f on S by

$$\|f\|_{L^q(S)} := \left(\int_S |f|^q dvol_S \right)^{1/q}.$$

For a given surface S in $\mathbb{R}^n$, the restriction problem asks to find all the inequalities of the form

$$\|\hat{f}\|_{L^q(S)} \leq C \|f\|_{L^p(\mathbb{R}^n)}. \tag{15.10}$$

Stein discovered several interesting things about this question. One important discovery is that the shape of the surface S matters. The unit sphere $S^{n-1} \subset \mathbb{R}^n$ obeys qualitatively different inequalities from a flat $(n-1)$-dimensional disk in $\mathbb{R}^n$. In particular the curvature of S plays an important role: when a surface is more curved, it obeys stronger inequalities.

At the time, it was counterintuitive to imagine there could be any interesting inequalities of this kind at all. Here is the issue. If $p = 1$, then it is easy to check that $\|\hat{f}\|_{L^\infty} \leq \|f\|_1$. This follows just from applying the triangle inequality to the definition of the Fourier transform:

$$|\hat{f}(\omega)| = \left| \int_{\mathbb{R}^n} f(x) e^{i\omega x} dx \right| \leq \int_{\mathbb{R}^n} |f(x)| dx.$$

Therefore, for any set S, we have the inequality $\|\hat{f}\|_{L^\infty(S)} \leq \|f\|_{L^1(\mathbb{R}^n)}$. If S is compact, like the unit sphere, we get $\|\hat{f}\|_{L^q(S)} \leq C(q,n)\|f\|_{L^1(\mathbb{R}^n)}$ for any q. The restriction question becomes interesting for $p > 1$. If $p > 1$, then $\|\hat{f}\|_{L^\infty}$ is not bounded by $\|f\|_{L^p}$. If S consists of a single point, ω_0, then $\|\hat{f}\|_{L^q(S)} = |\hat{f}(\omega_0)|$ for all q. So when S is a single point, and $p > 1$, then there are no inequalities of the form in Equation 15.10. Moreover, if S is an k-dimensional flat disk, with $1 \leq k \leq n-1$, and $p > 1$, then there are still no inequalities of the form in Equation 15.10. Based on these examples, it seemed plausible that there are no non-trivial inequalities of the form 15.10 at all. But it turns out that the situation is different when the surface S is curved, and then there are some inequalities with $p > 1$.

When the surface S is the unit sphere, Stein made a conjecture about all the restriction inequalities of the form in Equation 15.10, and he proved some non-trivial cases. The 2-dimensional case of this conjecture was proven by Fefferman in [**Fef2**]. It turns out that the restriction conjecture has a lot in common with the Bochner-Riesz conjecture that we discussed above. For instance, the restriction conjecture is connected to the Kakeya conjecture. Building on the work on the restriction problem by Stein and Fefferman, Carleson and Sjolin proved the 2-dimensional case of the Bochner-Riesz conjecture. In dimensions three and higher, all of these conjectures are wide open and look very difficult.

It is not clear how much polynomial arguments can contribute to understanding these problems. The problems are certainly not resolved, and there are serious obstacles to adapting the proof of finite field Kakeya to this setting, as we will discuss more below. On the other hand, polynomial methods have played a role in some recent results about this circle of problems. For example, the best currest estimate for the restriction problem in dimension three is based on polynomial partitioning [**Gu5**].

For an introduction to the restriction problem and the Kakeya problem, a good reference is Tao's lecture notes [**Ta2**]. For a discussion of the connection between the restriction problem and polynomial partitioning, see Lecture 3 of [**Gu6**].

We have now given some introduction to the connection between Kakeya-type problems and oscillatory integrals. In the rest of the chapter, we leave oscillatory integrals behind and focus on estimates for Kakeya-type problems.

15.6. Quantitative bounds for the Kakeya problem

In this section, we explore some quantitative lower bounds on the size of a Kakeya set $\cup_i T_i$. In the last section, we discussed how quantitative estimates about tubes are related to quantitative estimates about oscillatory integral operators. We have seen Besicovitch's example of a Kakeya set in two dimensions with $|\cup_i T_i| \lesssim \frac{\log\log N}{\log N} N^2$. In two dimensions, we will see that this example is essentially sharp. In higher dimensions, it is a major open problem whether the Besicovitch construction is essentially the best possible one. We will only prove some much weaker bounds. The methods that we discuss here are parallel to the methods we discussed for the finite field Kakeya problem in Section 3.1.

PROPOSITION 15.32. If $\{T_i\}$ is a Kakeya set of tubes in $\mathbb{R}^2$ (with width 1 and length N), then

$$|\cup_i T_i| \gtrsim (\log N)^{-1} N^2.$$

PROOF. The directions $v(T_i)$ are approximately evenly distributed on the unit circle. Therefore, we can number the tubes T_i so that

$$|v(T_i) - v(T_j)| \sim \frac{|i-j|}{N}.$$

In that case, the area of $T_i \cap T_j$ is bounded above

$$|T_i \cap T_j| \lesssim |v(T_i) - v(T_j)|^{-1} \lesssim N|i-j|^{-1}.$$

We can use this information to control $\int (\sum_i \chi_{T_i})^2$:

$$\int \left(\sum_{i=1}^N \chi_{T_i}\right)^2 = \sum_{i,j=1}^N \int \chi_{T_i}\chi_{T_j} \lesssim \sum_{i,j=1}^N N|i-j|^{-1} \sim (\log N) N^2.$$

On the other hand, if $K = \cup_i T_i$ is small, then $\int (\sum_i \chi_{T_i})^2$ is forced to be large. We can estimate this effect by using the Cauchy-Schwarz inequality.

$$N^2 \sim \int_K (\sum_i \chi_{T_i}) \cdot 1 \le |K|^{1/2} \left(\int (\sum_i \chi_{T_i})^2 \right)^{1/2} \lesssim |K|^{1/2} (N^2 \log N)^{1/2}.$$

Rearranging this inequality, we get the lower bound

$$|\cup_i T_i| = |K| \gtrsim N^2 (\log N)^{-1}.$$

□

Remark. This proof takes advantage of a connection between the size of $\cup_i T_i$ and the L^2 norm of $\sum \chi_{T_i}$. For a Kakeya set of tubes in n dimensions, we always have $\|\sum_i \chi_{T_i}\|_{L^1} \sim N^n$. If $|\cup_i T_i|$ is much smaller than N^n, then for any $p > 1$, $\int (\sum_i \chi_{T_i})^p$ is forced to be much larger than N^n. Partly for this reason, analysts are interested in estimating the L^p norms $\|\sum_i \chi_{T_i}\|_{L^p}$. Moreover, in the connections with linear operators that we described in the last section, these L^p norms are more relevant than the volume of $\cup_i T_i$.

One relevant example for the L^p estimates is the example when all the tubes of the Kakeya set are centered at the origin. In this case, $\sum_i \chi_{T_i}$ has size $\sim N^{n-1}$ on a

unit ball around the origin. For $p > \frac{n}{n-1}$ this peak on the unit ball is the dominant contribution to $\| \sum_i \chi_{T_i} \|_{L^p}$. Depending on the value of p either this example or the Besicovitch example give the largest known value of $\| \sum_i \chi_{T_i} \|_{L^p}$. The following conjecture implies that these examples are essentially the worst possible:

CONJECTURE 15.33. (L^p-Kakeya conjecture) Suppose that T_i is a Kakeya set of tubes in $\mathbb{R}^n$, and suppose that T_i^0 is a Kakeya set of tubes all centered at the origin. For any $p > \frac{n}{n-1}$,

$$\| \sum_i \chi_{T_i} \|_{L^p} \lesssim \| \sum_i \chi_{T_i^0} \|_{L^p}.$$

EXERCISE 15.10. Check that the L^p version of the Kakeya conjecture implies the Kakeya conjecture, Conjecture 15.21.

Using the same argument as in the proof of Proposition 15.32, one can check that in any dimension n, the L^p Kakeya conjecture holds for $p = 2$:

$$\| \sum_i \chi_{T_i} \|_{L^2} \lesssim \| \sum_i \chi_{T_i^0} \|_{L^2}. \tag{15.11}$$

In dimension $n \geq 3$, however, this estimate does not lead to sharp bounds for $| \cup_i T_i |$.

EXERCISE 15.11. Using Inequality 15.11, check that for all $n \geq 3$, if T_i is a Kakeya set of tubes, then $| \cup_i T_i | \gtrsim N^2$.

The bush argument is a simple argument that gives a much better bound for large values of n.

PROPOSITION 15.34. Suppose that $\{T_i\}$ is a Kakeya set of tubes in $\mathbb{R}^n$. Then $| \cup_i T_i | \gtrsim N^{\frac{n+1}{2}}$.

PROOF. Let $K = \cup_i T_i$. By the pigeon-hole principle, there must be a point $x \in K$ that lies in at least $\sum_i |T_i| |K|^{-1} \sim N^n |K|^{-1}$ different tubes T_i. Consider the union of all the tubes T_i containing the point x. This set of tubes is called the bush through x.

Far away from x, these tubes are morally disjoint from each other. Because the angle between any two tubes is $\gtrsim 1/N$, a point y outside of $B(x, N/4)$ can lie in $\lesssim 1$ different tubes in the bush through x. For each tube T_i in the bush through x, $|T_i \setminus B(x, N/4)| \gtrsim N$. Therefore, the union of the tubes in the bush has volume

$$\gtrsim \frac{N^n}{|K|} N.$$

Since the union of tubes in the bush is part of K, we get

$$|K| \gtrsim N^{n+1} |K|^{-1}.$$

Solving for $|K|$ yields $|K| \gtrsim N^{\frac{n+1}{2}}$.

□

The bush argument gives a better estimate than the L^2 argument for $n \geq 4$. On the other hand, the L^2 argument gives a better estimate when $n = 2$. (For $n = 3$, both arguments give $| \cup_i T_i | \gtrsim N^2$.) The hairbrush method, invented by

Tom Wolff, combines these two methods in a clever way to give a better estimate. It leads to the bound

$$|\sum_i T_i| \gtrsim N^{\frac{n+2}{2}}.$$

We described the finite field version of the hairbrush argument in Section 2.4. It could be an interesting project for the reader to try to generalize this discussion to tubes in $\mathbb{R}^n$.

Mathematicians have tried hard to improve these estimates, and it seems to be very difficult to achieve sharp bounds. In the late 90's, Bourgain introduced a new approach to this problem using combinatorial number theory. This new approach led to much better bounds for large n. It was developed further by Katz and Tao. For large n, Katz and Tao [**KatTar**] prove that $|\cup_i T_i| \gtrsim N^{\alpha n}$ for $\alpha = .59...$. We don't discuss the number theory approach to Kakeya here. For an introduction to these ideas, see [**Lab**], [**Ta1**], or [**D2**].

In three dimensions, the hairbrush argument gives the bound $|\cup T_i| \gtrsim N^{\frac{5}{2}}$. This bound has been quite difficult to improve. Combining the combinatorial number theory tools with other interesting ideas, Katz, Laba, and Tao, under a small extra assumption about the tubes T_i, improved the bound to $N^{\frac{5}{2}+\varepsilon}$ for some small $\varepsilon > 0$. This result was published in the Annals of Math. It remains the best result in three dimensions at the present.

15.7. The polynomial method and the Kakeya problem

The finite field Kakeya problem was introduced by Wolff in the late 90's. Before Dvir's work, essentially the same estimates were known for the original Kakeya problem and the finite field Kakeya problem, although there were small technical differences in some of the proofs. It came as a real shock to the community when Dvir completely solved the finite field Kakeya problem in a couple pages [**D**].

It is not yet clear how much the polynomial method can say about problems in harmonic analysis like the Kakeya problem or the Bochner-Riesz conjecture. It's not clear whether this short proof over finite fields is a crucial clue or a red herring. People have tried hard to adapt the proof to Euclidean space, and there are some serious difficulties. On the other hand, there have been some small successes where the polynomial method has led to harmonic analysis estimates that are slightly stronger than what we can prove without polynomials.

In this section, we describe some of the difficulties of adapting the polynomial method to the Kakeya problem for tubes, but we will also see that polynomials do tell us something about Kakeya sets.

To get started, let us make a high-level sketch of the proof of the finite field Kakeya conjecture from Section 2.4. Suppose that $K \subset \mathbb{F}_q^n$ is a Kakeya set. We consider a polynomial P of minimal degree that vanishes on K. First we estimate the degree of P. If $|K|$ is much smaller than q^n, then we see that $\operatorname{Deg} P \leq q/2$. Now by the vanishing lemma, if P vanishes on a line $l \subset \mathbb{F}_q^n$, then P must also vanish at the point at infinity corresponding to l. But then P would vanish at too many places, giving a contradiction.

Suppose now that $\{T_i\}$ is a Kakeya set of tubes in $\mathbb{R}^n$. Each tube is a cylinder of radius 1 and length N, and the angle between any two tubes is $\gtrsim 1/N$, and the number of tubes is $\sim N^{n-1}$. We let $K := \cup_i T_i$. Let us suppose that $|K| \sim N^{n-\gamma}$ for some $\gamma > 0$.

How can we imitate the proof of the finite field Kakeya conjecture? There is no (non-zero) polynomial P that vanishes on the entire set K, because K contains an open ball. If we take a finite set of points in K, we could find a polynomial that vanishes on that set of points. For example, we might consider a unit lattice and look at the set of all the unit cubes in the lattice that intersect K. We let $\mathcal{Q}(K)$ denote this set of cubes. The number of such unit cubes is $\sim |K| \sim N^{n-\gamma}$. We could let P be a minimal degree polynomial that vanishes at the center of each of these cubes. By parameter counting, we get $\operatorname{Deg} P \lesssim N^{n-\gamma}$. Each tube would contain $\sim N$ points where P vanishes, and N is much larger than $\operatorname{Deg} P$. However, it is not clear how to make use of this. For almost all the tubes T_i, these $\sim N$ points in $Z(P)$ do not perfectly line up along a line, so we cannot apply the vanishing lemma.

The key difficulty in adapting the proof of finite field Kakeya is in the step with the vanishing lemma. If a polynomial P vanishes at more than $\operatorname{Deg} P$ points on a line l, then P must vanish on the whole line l, including points very far away from the initial points. This is a very strong result, and it plays a crucial role in the polynomial method: a polynomial that vanishes at some points is forced to vanish somewhere else. In the situation above, we know that P vanishes at $\sim N$ points of T_i, and these N points are roughly spaced in a row along T_i with a distance of ~ 1 between consecutive points. Recall that $v(T_i)$ denotes a unit vector in the direction of T_i. Suppose that T_i^+ is defined to be the translation of T_i by $Nv(T_i)$. When I first started working on the problem, I was hoping to prove that P vanishes at many evenly spaced points along T_i^+. But this doesn't have to happen. The variety $Z(P)$ can hug a tube T_i for a long way and then turn sharply away. This is illustrated in Figure 15.3. For an example of an algebraic curve that makes a sharp turn, consider the polynomial curve in two variables defined by $y = x^D$. For $0 \le x \le 1 - D^{-1/2}$, this curve closely hugs the x-axis. But around $x = 1$, the curve turns sharply and hugs the line $x = 1$ instead. Just because the polynomial P vanishes at many points along T_i, it doesn't seem to force P to vanish anywhere far from T_i. So a crucial part of the polynomial method breaks down here.

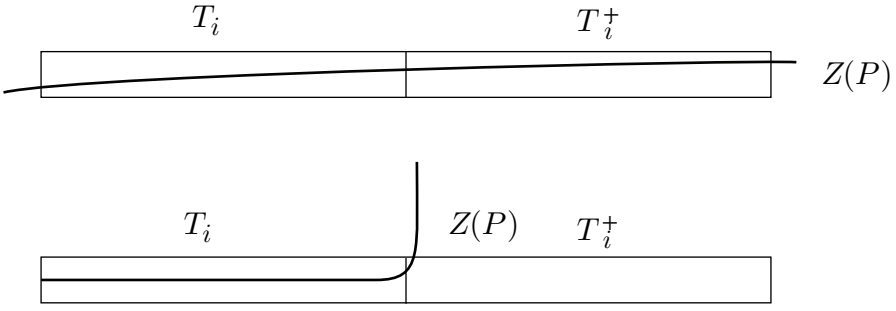

FIGURE 15.3. The top scenario works well for polynomial methods, but the bottom scenario can happen too.

It is easy to construct an algebraic surface $Z(P)$ that vanishes at many evenly spaced points along one tube T_i, but barely intersects the translated tube T_i^+. It's not so easy to imagine an algebraic surface $Z(P)$ that does the same to all the tubes T_i in a hypothetical Kakeya set. It may be possible to attack the Kakeya problem

by proving that an algebraic surface $Z(P)$ cannot bend too sharply in too many places, in terms of the degree of P.

Even simple problems about sharp bending of algebraic surfaces turn out to be rather hard. For example, when I taught a course on the polynomial question, I raised the question how sharply a degree D algebraic curve in the plane can bend. We can make this precise as follows. We say that a polynomial $P \in \mathrm{Poly}_D(\mathbb{R}^2)$ makes an ε-sharp right-angled turn at zero if

- $P(x_1, x_2) > 0$ if $(x_1, x_2) \in [\varepsilon, 1]^2$,
- $P(x_1, x_2) < 0$ if $(x_1, x_2) \in [-1, -\varepsilon] \times [-1, 1]$ or $(x_1, x_2) \in [-1, 1] \times [-1, -\varepsilon]$.

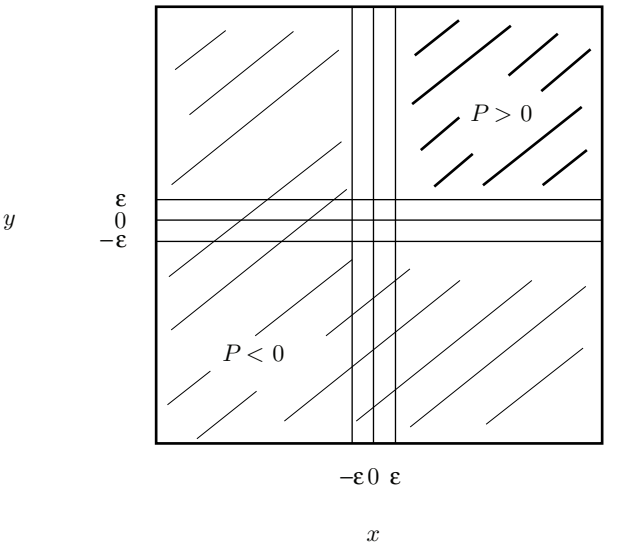

FIGURE 15.4

Define $\varepsilon(D)$ to be the infimal value of ε so that some polynomial $P \in \mathrm{Poly}_D(\mathbb{R}^2)$ makes an ε-sharp right-angled turn at 0. It turns out to be hard just to estimate the asymptotic behavior of $\varepsilon(D)$ as $D \to \infty$. If D is even, then the function $P(x_1, x_2) = 1 - (1 - x_1)^D - (1 - x_2)^D$ makes an ε-sharp turn for $\varepsilon \sim 1/D$. I conjectured that this example is asymptotically sharp in the sense that $\varepsilon(D) \sim 1/D$. In [**Z**], Zhang constructed examples showing that $\varepsilon(D) \lesssim e^{-cD}$ for some constant $c > 0$, showing that my conjecture was badly wrong. In the other direction, he proved that $\varepsilon(D) \gtrsim e^{-CD^2}$ for some constant C. These bounds give a fairly precise picture of the asymptotics of $\varepsilon(D)$. Zhang's example shows that algebraic curves can bend surprisingly fast, and it tends to make me pessimistic about controlling the behavior of $Z(P)$ on the shifted tubes T_i^+.

One part of the difficulty in working with tubes is that a finite set of points is a very thin subset of the n-dimensional tube T_i. We can get some additional leverage by applying the polynomial ham sandwich theorem. In the rest of this section, we explore in a heuristic way what the polynomial ham sandwich theorem can tell us about Kakeya sets. In the next section, we will use these ideas prove a precise theorem.

Using the polynomial ham sandwich theorem, we can find a non-zero polynomial P so that $Z(P)$ bisects each cube $Q \in \mathcal{Q}(K)$, and with

$$\operatorname{Deg} P \lesssim |\mathcal{Q}(K)|^{1/n} \lesssim N^{1-\frac{\gamma}{n}}.$$

If $Z(P)$ bisects a unit cube Q, then $\operatorname{Vol}_{n-1} Z(P) \cap Q \gtrsim 1$, and so we know that $Z(P) \cap Q$ has a substantial surface area instead of just knowing that $Z(P) \cap Q$ contains the point at the center of Q.

Consider one of the tubes, T_i. We let D_i be an orthogonal cross-section of T_i, and for every $x \in D_i$ we let ℓ_x be the line through x parallel to T_i. For almost every choice of $x \in D_i$, we have

$$|\ell_x \cap Z(P)| \leq \operatorname{Deg}(P) \lesssim N^{1-\frac{\gamma}{n}}. \tag{15.12}$$

On the other hand, the tube T_i contains $\sim N$ cubes of $\mathcal{Q}(K)$, and $Z(P)$ bisects each of them. Let $\mathcal{Q}(T_i) \subset \mathcal{Q}(K)$ be the set of cubes of $\mathcal{Q}(K)$ that intersect T_i. They are disjoint, so we get an estimate for the average: for almost every $x \in D_i$,

$$\operatorname{Avg}_{Q \in \mathcal{Q}(T_i)} |\ell_x \cap Z(P) \cap Q| \lesssim N^{-1} \operatorname{Deg}(P) \lesssim N^{-\frac{\gamma}{n}}.$$

Since this holds for almost every $x \in D_i$, it also holds when we average over $x \in D_i$. So we get

$$\operatorname{Avg}_{Q \in \mathcal{Q}(T_i), x \in D_i} |\ell_x \cap Z(P) \cap Q| \lesssim N^{-\frac{\gamma}{n}}.$$

For a typical cube $Q \in \mathcal{Q}(T_i)$, we know that $Z(P)$ bisects Q, and so $\operatorname{Vol}_{n-1} Z(P) \cap Q \gtrsim 1$, and yet

$$\operatorname{Avg}_{x \in D_i} |Z(P) \cap Q \cap \ell_x| \lesssim N^{-\gamma/n} \ll 1.$$

This is only possible if the surface $Z(P) \cap Q$ is approximately parallel to the tube T_i! Figure 15.5 illustrates what we have learned about how $Z(P)$ intersects a tube T.

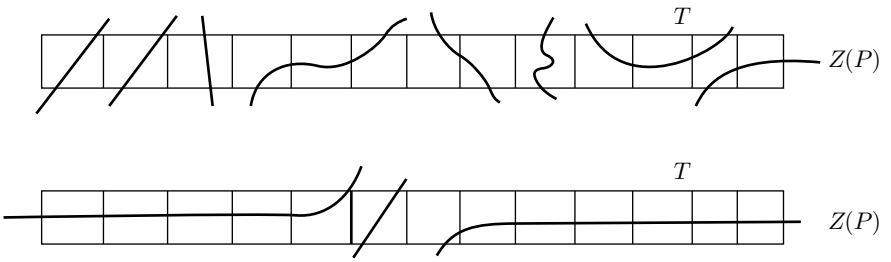

FIGURE 15.5. The top picture is an impossible picture of how $Z(P)$ intersects T. The bottom picture is a more realistic picture of how $Z(P)$ may intersect T. Note here that $Z(P)$ is usually almost tangent to the direction of T.

This observation has interesting consequences for the structure of a Kakeya set. We noticed that for a typical cube $Q \in \mathcal{Q}(T_i)$, the tube T_i is approximately parallel to the surface $Z(P) \cap Q$. But there are many different tubes T_j of our Kakeya set that intersect a typical cube Q! By the argument above, almost all of these tubes are approximately parallel to the surface $Z(P) \cap Q$. Therefore, there must be a hyperplane $\pi(Q)$, and the tubes T_j intersecting Q must usually be almost tangent to $\pi(Q)$.

This is a surprising structure, called planiness. Without any experience, we might expect that the different tubes T_j intersecting a cube Q would point in a complicated set of directions on the unit sphere. Suprisingly, they need to concentrate near to a plane.

Planiness was first discovered by Katz, Laba, and Tao, in the paper [**KLT**]. Planiness was one of the observations/tools that allowed them to prove that a Kakeya set of tubes in $\mathbb{R}^3$ (with mild additional hypotheses) has volume at least $N^{2.5+\epsilon}$. Later, Bennett, Carbery, and Tao proved stronger and more general planiness estimates in the paper [**BCT**]. The polynomial method gives a third approach to planiness, explained in [**Gu3**].

To end this section, we return to the vanishing lemma in the context of tubes. Recall that the original vanishing lemma says that if a polynomial P vanishes at $> \operatorname{Deg} P$ points on a line l, then P vanishes on the whole line. This result does not seem to generalize from lines to tubes in any nice way. But let us consider a weak corollary of the vanishing lemma.

COROLLARY 15.35. Suppose that P vanishes at some points $x_1, ..., x_M$ along a line l. Let v be a vector parallel to l. If $M > \operatorname{Deg} P$, then $\nabla_v P(x_i) = 0$ for $i = 1, ..., M$. In other words, $Z(P)$ is tangent to l at the points $x_1, ..., x_M$.

The discussion above and the pictures in Figure 15.5 show that this weak corollary of the vanishing lemma does generalize to tubes in a fairly nice way.

This weak vanishing lemma, Corollary 15.35, is all that we actually used in the proof of the joints theorem in Section 2.5. In the next section, we will use these ideas to prove a version of the joints theorem for tubes.

15.8. A joints theorem for tubes

During the book, we have met many theorems about the incidence patterns of lines in space. Each of these questions can be adapted to a question about long thin tubes instead of lines. In many cases, the tube version is wide open. But for the joints problem, there is a version for tubes that has a nice proof using the ideas from the last section.

THEOREM 15.36. ([**BCT**], [**Gu3**]) Suppose that $\ell_{j,a}$ are lines in $\mathbb{R}^n$ for $1 \leq j \leq n$ and $1 \leq a \leq A$, where each line $\ell_{j,a}$ makes an angle of at most $(100n)^{-1}$ with the x_j-axis. Let $T_{j,a}$ be the infinite cylinder with radius 1 centered on $\ell_{j,a}$.

Let I be the set of points which lie in one cylinder for each value of $j = 1...n$. In equations

$$I := \cap_{j=1}^{n} (\cup_{a=1}^{A} T_{j,a}).$$

Then the volume of I is $\lesssim A^{\frac{n}{n-1}}$.

This theorem is analogous to the joints theorem. Each point of I is like a "joint" of the tubes $T_{j,a}$. So the volume of I is analogous to the number of joints. If a point x lies in one line for each value of $j = 1...n$, then x is a joint for the set of lines $\ell_{j,a}$. The joints theorem says that the number of joints is $\lesssim A^{\frac{n}{n-1}}$. This theorem says that the volume of I obeys the same estimate, generalizing the joints theorem to tubes.

If the tubes $T_{j,a}$ are parallel to the x_j-axis, then this estimate follows from the Loomis-Whitney inequality, Theorem 15.1. The projection of I to any coordinate hyperplane lies in the union of A unit balls, and so for each $j = 1, ..., n$ we get

$|\pi_j(I)| \lesssim A$. Theorem 15.1 then gives $|I| \lesssim A^{\frac{n}{n-1}}$. So Theorem 15.36 can be thought of as a generalization of the Loomis-Whitney inequality where the tubes are allowed to tilt a few degrees.

Bennett, Carbery, and Tao proved a very slightly weaker version of Theorem 15.36 in [**BCT**]. They showed that for any $\varepsilon > 0$, $|I| \leq C(n,\varepsilon)A^{\frac{n}{n-1}+\varepsilon}$. Their proof was based on a monotonicity formula for the heat equation, and it involved looking at the problem at multiple scales.

We will prove Theorem 15.36 using the polynomial ham sandwich theorem. This approach gives the estimate $I \leq C(n)A^{\frac{n}{n-1}}$, which is sharp up to a constant factor. This estimate is only very slightly stronger than what we can prove without polynomials, but it still indicates that polynomials have something interesting to contribute to studying the intersection patterns of tubes in Euclidean space.

The proof also involves the idea of the directed volume of a surface. Suppose S is a smooth hypersurface in $\mathbb{R}^n$. For each $x \in S$, let $N(x)$ denote a unit normal vector to the surface S at the point x. If v is a unit vector, we define the directed volume of S perpendicular to V by the formula

$$V_S(v) := \int_S |N \cdot v| dvol_S.$$

Notice that if the tangent plane of S is perpendicular to v, we have $|N \cdot v| = 1$, and if the tangent plane contains v, we have $|N \cdot v| = 0$. For example, we consider the directed volume of the unit circle in the direction $v = (0,1)$. The directed volume of an arc of the upper semi-circle in direction v is exactly the change in the x-coordinate over the arc. Therefore, the directed volume of the whole upper semi circle is 2, and the directed volume of the whole circle is 4.

The computation for the circle generalizes as follows. Let π be the orthogonal projection from $\mathbb{R}^n$ to $v^\perp \subset \mathbb{R}^n$.

LEMMA 15.37. $V_S(v) = \int_{v^\perp} |S \cap \pi^{-1}(y)| dvol(y)$.

As a corollary, we can immediately estimate the directed volume of a degree d variety in a cylinder T.

LEMMA 15.38. (Cylinder estimate) Let T be an infinite cylinder in $\mathbb{R}^n$ of radius r. Let v be a unit vector parallel to the axis of T. Let $Z(P)$ be the vanishing set of a polynomial P.

Then $V_{Z(P)\cap T}(v) \lesssim r^{n-1} deg(P)$.

PROOF. Let π be the projection from T to the cross-section $v^\perp \cap T$. This cross-section is just an (n-1)-dimensional ball of radius r. For almost every y in this ball, $|\pi^{-1}(y) \cap Z(P)| \leq \mathrm{Deg}(P)$. By the last lemma, $V_{Z(P)\cap T}(v)$ is bounded by $\mathrm{Deg}(P)$ times the volume of the cross-section, which is $\sim r^{n-1}$. □

LEMMA 15.39. If S is a hypersurface in $\mathbb{R}^n$, and $v_1, ..., v_n$ are unit vectors and the angle from v_j to the x_j-axis is $\leq (100n)^{-1}$, then $Vol_{n-1} S \leq 2\sum_j V_S(v_j)$.

PROOF. At a given point of S with normal vector N, we have to prove that $\sum_j |N \cdot v_j| \geq 1/2$. If e_j are the coordinate vectors, then for any unit vector N we have

$$\sum_j |N \cdot e_j| \geq \sum_j |N \cdot e_j|^2 = 1.$$

The vectors v_j are very close to e_j, and so we get

$$\sum_j |N \cdot v_j| \geq \sum_j |N \cdot e_j| - \sum_j |v_j - e_j| \geq 1 - \frac{1}{100} \geq 1/2.$$

Now

$$\sum_j V_S(v_j) = \int_S \left(\sum_j |N(x) \cdot v_j| \right) dvol_S \geq (1/2) \operatorname{Vol} S.$$

□

Now we can prove Theorem 15.36.

PROOF. Consider the unit cubical lattice. Let $Q_1, ..., Q_V$ be all the unit cubes in the lattice which intersect the set I. We will prove $V \lesssim A^{\frac{n}{n-1}}$.

Let P be a non-zero polynomial so that $Z(P)$ bisects each cube $Q_1, ..., Q_V$ and $\operatorname{Deg} P \lesssim V^{1/n}$. This bisection requires a certain amount of area, therefore:

$$Vol_{n-1} Z(P) \cap Q_i \gtrsim 1.$$

Since Q_i intersects I, Q_i must intersect $\cup_a T_{j,a}$ for each j. For each cube Q_i, for each j, we let $T_j(Q_i)$ be one of the tubes $T_{j,a}$ that intersects Q_i. Let $v_{j,i}$ be the direction of the tube $T_j(Q_i)$. By Lemma 15.39, we get

$$\sum_{j=1}^{n} V_{Z(P) \cap Q_i}(v_{j,i}) \gtrsim Vol_{n-1} Z(P) \cap Q_i \gtrsim 1.$$

For each cube Q_i, choose one direction $j(Q_i)$ so that $V_{Z(P) \cap Q_i}(v_{j,i}) \gtrsim 1$. Then assign the cube Q_i to the tube $T_j(Q_i)$. Each cube is now assigned to a tube. We have V cubes and nA tubes, so one of the tubes has $\gtrsim V/A$ cubes assigned to it. Let T be this tube, and let v be its direction. We have $\gtrsim V/A$ cubes Q_i obeying the following conditions:

- The cube Q_i intersects T.
- $V_{Z(P) \cap Q_i}(v) \gtrsim 1$.

Let $\tilde{T}$ be a wider cylinder with radius $2n$ and with the same central axis as T. If Q_i intersects T, then it lies in $\tilde{T}$. Therefore, we have

$$V/A \lesssim V_{Z(P) \cap \tilde{T}}(v).$$

On the other hand, the cylinder estimate, Lemma 15.38, gives

$$V_{Z(P) \cap \tilde{T}}(v) \lesssim V^{1/n}.$$

Combining these inequalities we get $V/A \lesssim V^{1/n}$, and rearranging we get $V \lesssim A^{\frac{n}{n-1}}$. □

15.9. Hermitian varieties

In this section, we describe an example of a set of complex tubes in $\mathbb{C}^3$ with a remarkable intersection pattern. A complex line segment in $\mathbb{C}^n$ of length L is defined to be the intersection of a complex line $l \subset \mathbb{C}^n$ with a ball $B(z, L)$ for some point $z \in l$. A complex tube of radius δ and length 1 is defined to be the δ-neighborhood of a complex line segment of length 1. Geometrically, a complex tube of radius δ and length 1 in $\mathbb{C}^n$ looks approximately like a pancake of the form $B^{2n-2}(\delta) \times B^2(1)$.

We will consider a set of complex tubes $T_i \subset \mathbb{C}^3$ of radius δ and length 1. We are interested in sets of tubes that overlap a great deal in the sense that $|\cup_i T_i|$ is much smaller than $\sum_i |T_i|$. One boring way that this could happen is that all the tubes T_i could be miniscule perturbations of a single tube. More generally, many tubes T_i could pack into a tube of radius w and length 1 for some $\delta \leq w \leq 1$. We could easily get compression if such a tube $T(w)$ contained a set of our tubes T_i with

$$\sum_{T_i \subset T(w)} |T_i| \ll |T(w)|.$$

The volume of a tube of radius w and length 1 in $\mathbb{C}^3$ is $\sim w^4$. Therefore, we will be interested in sets of tubes T_i so that

(15.13) Any tube of radius w and length 1 contains $\lesssim (w/\delta)^4$ tubes T_i.

Another way that the tubes T_i could overlap a lot is that many tubes could cluster into a small neighborhood of a plane. If π is a plane, then the volume of $N_w(\pi) \cap B(1)$ is $\sim w^2$. Therefore, we will be interested in sets of tubes T_i so that for any complex 2-plane $\pi \subset \mathbb{C}^3$, and any $\delta < w < 1$,

(15.14) the number of tubes T_i in the slab $N_w(\pi) \cap B(1)$ is $\lesssim \delta^{-2}(w/\delta)^2$.

Under these conditions, it is not easy to imagine an example where $|\cup_i T_i| \ll \sum_i |T_i|$, but it can happen.

THEOREM 15.40. ([**KLT**]) There exists a set of δ^{-4} complex tubes T_i obeying Conditions 15.13 and 15.14, with

$$|\cup_i T_i| \lesssim \delta \sim \delta \sum_i |T_i|.$$

The union $\cup_i T_i$ contains $\sim \delta^{-5}$ disjoint δ cubes, each cube lies in $\sim \delta^{-1}$ tubes T_i, and each T_i intersects $\sim \delta^{-2}$ cubes.

The example in [**KLT**] is based on the Heisenberg variety. Steve Kleiman explained to me that there are similar examples based on Hermitian varieties, which were studied in algebraic geometry by Bose and Chakravarti [**BC**]). In Section 3.2, we studied Hermitian varieties in finite fields, and in this section, we study the analogous varieties over $\mathbb{C}$.

It is interesting to contrast this theorem with our results about the intersection patterns of complex lines in $\mathbb{C}^3$. In the exercises, we showed the following estimate about complex lines – see Exercise 11.4 for an approach using flat points and Exercise 13.4 for an approach using ruled surfaces.

THEOREM 15.41. Suppose that $\mathfrak{L}$ is a set of N^2 lines in $\mathbb{C}^3$ with at most N lines in any plane. Suppose that $\mathcal{S}$ is a set of points in $\mathbb{C}^3$ so that each line of $\mathfrak{L}$ contains at least N points of $\mathcal{S}$. Then $|\mathcal{S}| \gtrsim N^3$.

Comparing Theorem 15.40 and Theorem 15.41, we see that the incidence geometry of thin tubes in $\mathbb{C}^3$ is very different from the incidence geometry of lines in $\mathbb{C}^3$. (In this comparison, δ^{-2} plays the role of N.)

This example shows that it is complicated to try to adapt theorems about lines to theorems about thin tubes. Some results generalize from lines to thin tubes and some do not. In Section 15.8, we saw that the joints theorem generalizes to thin tubes in a nice way. But the Hermitian variety shows that Theorem 15.41 fails to generalize to thin tubes.

For many other problems about the incidence geometry of lines, we don't have a good understanding of what happens for thin tubes. The Szemerédi-Trotter theorem is an important example, which was studied by Wolff [**Wo2**], Katz-Tao [**KT2**], and Bourgain [**Bou**] . Building on the previous work, the paper [**Bou**] proves a very interesting estimate about the Szemerédi-Trotter problem for thin tubes, but this estimate is far from sharp, and we do not have anything like a complete understanding.

Theorem 15.40 is also relevant for the Kakeya problem. We can make a complex analogue of the Kakeya problem in the following way. We say that the angle between two 1-dimensional subspaces $V_1, V_2 \subset \mathbb{C}^3$ is at most δ if $N_\delta V_1 \supset V_2 \cap B(1)$. (Exercise: it would be equivalent to say that $N_\delta V_2 \supset V_1 \cap B(1)$.) We say that the angle between two tubes is $\geq \delta$ if the angle between the subspaces tangent to the central axes of the tubes is $\geq \delta$. A Kakeya set of tubes of radius δ and length 1 in $\mathbb{C}^n$ is a set of $\delta^{-(2n-2)}$ tubes where the angle between any two tubes is $\gtrsim \delta$. It is straightforward to check that a Kakeya set of tubes in $\mathbb{C}^3$ obeys the conditions 15.13 and 15.14, and there are similar conditions in any dimension n. The complex Kakeya conjecture says that for any Kakeya set of tubes in $\mathbb{C}^n$, and any $\varepsilon > 0$, $|\cup_i T_i| \geq c(n,\varepsilon)\delta^\varepsilon$.

The configuration of tubes coming from the Hermitian variety or the Heisenberg group is not a Kakeya set of tubes. It contains many parallel tubes. But this set of tubes does obey the interesting conditions 15.13 and 15.14. Many arguments about Kakeya sets in $\mathbb{R}^3$ only really used these two conditions: in particular, the hairbrush argument only used these two conditions. In the paper [**KLT**], Katz, Laba, and Tao proved a Kakeya estimate in $\mathbb{R}^3$ that improved on the bound from the hairbrush argument, although only by a small margin. In order to improve the hairbrush bound, they had to use some hypothesis that is obeyed by a Kakeya set of tubes in $\mathbb{R}^3$ but not obeyed by the tubes from Theorem 15.40. This means either using that the tubes are real instead of complex, and/or using that they point in different directions instead of just using Conditions 15.13 and 15.14. The paper [**KLT**] figured out how to exploit this information to get improved estimates for the size of a Kakeya set.

Now we turn to the proof of Theorem 15.40. We will prove a slightly weaker result, where the estimate in Condition 15.14 is weaker by a factor $(\log \delta^{-1})$. The construction is based on an interesting variety called the Hermitian variety. It is the complex analogue of the Hermitian varieties over finite fields that we discussed in Section 3.2. The most direct analogue of the Hermitian variety over finite fields would be given by the equation $|z_1|^2 + |z_2|^2 + |z_3|^2 = 1$. This variety is just a 5-dimensional (real) sphere in $\mathbb{C}^3$. It is compact and so it does not contain any lines. By modifying the signs on the left-hand side, we get an interesting variety that contains many lines. We define the variety H as follows:

$$H := \{(z_1, z_2, z_3) \in \mathbb{C}^3 \text{ so that } |z_1|^2 + |z_2|^2 - |z_3|^2 = 1\}.$$

(The set H is not a complex algebraic variety in $\mathbb{C}^3$, but we can think of it as a real algebraic hypersurface in $\mathbb{R}^6$.)

The set H has a lot symmetries and it contains a lot of lines. First we describe the symmetries. We define a (non-standard) Hermitian inner product on $\mathbb{C}^3$ by

$$(v, w)_H := v_1\bar{w}_1 + v_2\bar{w}_2 - v_3\bar{w}_3.$$

Notice that $H = \{z \in \mathbb{C}^3 | (z, z)_H = 1\}$. The group $U(2, 1) \subset GL(3, \mathbb{C})$ is defined to be the set of complex-linear isomorphisms of $\mathbb{C}^3$ that preserve this inner product. In other words

$$U(2, 1) := \{M \in GL(3, \mathbb{C}) | (Mv, Mw)_H = (v, w)_H \text{ for all } v, w \in \mathbb{C}^3\}.$$

The group $U(2, 1)$ acts on H and we will see in the next exercise that it acts transitively. This shows that H has a lot of symmetry.

The group $U(2, 1)$ has a lot in common with the standard unitary group $U(3)$ corresponding to the standard Hermitian inner product on $\mathbb{C}^3$. For instance, there are many orthonormal bases, described in the following exercise.

Exercise 15.12. Suppose that $v \in \mathbb{C}^3$ with $(v, v)_H = 1$. Then there is a basis $v_1 = v$, v_2, v_3 of $\mathbb{C}^3$ with $(v_i, v_j)_H = \delta_{ij}$. (In other words, $(v_i, v_j)_H$ is equal to 1 if $i = j$ and equal to zero if $i \neq j$.)

As a corollary, $U(2, 1)$ acts transitively on H. For any $v \in H$, there is a matrix $M \in U(2, 1)$ so that $M(1, 0, 0) = v$. To find M, we let v_1, v_2, v_3 be the basis defined in the first part of the problem, and we let M be the matrix with columns v_1, v_2, and v_3. We then check that $M \in U(2, 1)$ and $M(1, 0, 0) = v$.

Next we study the complex lines in H. Since $U(2, 1)$ acts transitively on H, it suffices to study the complex lines in H through the point $(1, 0, 0)$. Notice that $H \cap \{z_1 = 1\}$ is the set defined by the following equations:

$$z_1 = 1; |z_2|^2 = |z_3|^2.$$

This set contains many complex lines: all the complex lines of the form $z_1 = 1, z_3 = \alpha z_2$, for any $\alpha \in \mathbb{C}$ with $|\alpha| = 1$.

Conversely, these lines are all of the lines through $(1, 0, 0)$ contained in H. Suppose that l is a complex line with $(1, 0, 0) \subset l \subset H$. We identify $\mathbb{C}^3$ with $\mathbb{R}^6$ and use coordinates x_j, y_j on $\mathbb{R}^6$ with $z_j = x_j + iy_j$. Now we can think of l as a real 2-plane in the real 5-manifold H. The (real) tangent space to H at $(1, 0, 0)$ is the space $x_1 = 1$. Therefore, l lies in the real hyperplane $x_1 = 1$. But since l is a complex line, we can parametrize it by a map $\mathbb{C} \to \mathbb{C}^3$ with $t \mapsto (a_1 t + b_1, ..., a_3 t + b_3)$. We see that the real part of $a_1 t + b_1 = 1$ for all $t \in \mathbb{C}$. But this is only possible if $a_1 = 0$ and $b_1 = 1$. So we conclude that the line l must lie in the complex 2-plane $z_1 = 1$.

This finishes our description of the lines in H through $(1, 0, 0)$. We restate it as a lemma.

Lemma 15.42. The set of complex lines l with $(1, 0, 0) \in l \subset H$ is parametrized by the unit circle $S^1 \subset \mathbb{C}$. For each $\alpha \in S^1$, we define a line l_α by

$$z_1 = 1; z_2 = \alpha z_3.$$

The set of complex lines l with $(1, 0, 0) \in l \subset H$ is $\{l_\alpha\}_{\alpha \in S^1}$.

We now construct the set of tubes T_i by thickening some of the lines in H. First, we choose a δ-separated set of points $a \in S^1$. The number of points a is $\sim \delta^{-1}$. We consider the lines l_a defined above. Next we use the action of $U(2, 1)$ to translate these lines so that they go through other points of H. We let z_b be a set of δ-separated points in $H \cap B(2)$. The number of points z_b is $\sim \delta^{-5}$. For

each z_b, we choose a matrix $M_b \in U(2,1)$ so that $M_b(1,0,0) = z_b$. Then we define $l_{ab} = M_b(l_a)$. The line l_{ab} lies in H and contains the point z_b. We let T_{ab} be the δ-neighborhood of $l_{ab} \cap B(2)$.

We have now defined $\sim \delta^{-6}$ tubes T_{ab}. Not all these tubes are really distinct. Define two tubes to be equivalent if each one is contained in the δ-neighborhood of the other. We keep only one tube from each equivalence class, and we let T_i be the resulting set of tubes.

We can now estimate the number of tubes T_i. For each point z_b, the angle between any two tubes T_{ab} is $\gtrsim \delta$, and therefore, there are $\gtrsim \delta^{-1}$ inequivalent tubes T_{ab} though z_b. Because of the description of the lines in H in Lemma 15.42, there are at most $\lesssim \delta^{-1}$ inequivalent tubes T_i that intersect $(1,0,0)$. Because of the symmetries of H, there are at most $\lesssim \delta^{-1}$ inequivalent tubes T_i that pass through any point in $H \cap B(2)$. Therefore, there are $\sim \delta^{-1}$ tubes T_i that pass through each point z_b. A little more generally, we can cover $H \cap B(2)$ with $\sim \delta^{-5}$ δ-cubes Q_b (one around each point $z(b)$), and each δ-cube intersects $\sim \delta^{-1}$ tubes T_i. Since each tube T_i intersects $\sim \delta^{-2}$ of these cubes, we see that the number of tubes T_i is $\sim \delta^{-4}$.

To complete the proof of Theorem 15.40, we have to check that the tubes T_i obey Conditions 15.13 and 15.14 – that they don't cluster into a thicker tube or into a planar slab. We begin by checking Condition 15.13.

Let $T(w)$ be a tube of radius w and length 1, for some $\delta < w < 1$. First we consider $T(w) \cap N_\delta(H)$. Since H is a real hypersurface of low degree, this set has volume at most δw^3, and so it can be covered by $\sim \delta^{-2}(w/\delta)^3$ δ-cubes Q_b. Each of these cubes lies in $\sim \delta^{-1}$ tubes from our set T_i. These δ^{-1} tubes point in a range of directions, described in Lemma 15.42. The number of tubes T_i through a cube Q_b that make an angle $\lesssim w$ with a given fixed direction is $\lesssim (w/\delta)$. Therefore, each cube Q_b lies in $\lesssim w/\delta$ tubes $T_i \subset T(w)$. Since each tube $T_i \subset T(w)$ contains $\sim \delta^{-2}$ cubes $Q_b \subset T(w)$, we see that the number of tubes $T_i \subset T(w)$ is at most

$$\delta^2 \cdot \delta^{-2}(w/\delta)^3 \cdot (w/\delta) \sim (w/\delta)^4.$$

This proves Condition 15.13.

Next we discuss Condition 15.14: the tubes T_i do not cluster in a planar slab. If π is a (complex) 2-plane in $\mathbb{C}^3$ then the heart of the matter is an estimate for the volume of $N_\delta(\pi) \cap N_\delta(H) \cap B(2)$. We will prove the following bound.

$$|N_\delta(\pi) \cap N_\delta(H) \cap B(2)| \lesssim \delta^3 \log(\delta^{-1}). \tag{15.15}$$

Let us take a moment to process the expression $\delta^3 \log(\delta^{-1})$. The volume of $N_\delta(\pi) \cap B(2)$ is $\sim \delta^2$. If H were a real 5-plane transverse to the real 4-plane π, then the intersection would have volume $\sim \delta^3$. On the other hand, if H were a real 5-plane containing π, then the intersection would be all of $N_\delta(\pi) \cap B(2)$, with volume around δ^2. The factor $\log(\delta^{-1})$ is not really necessary in Equation 15.15, but the factor is small and it allows a simpler proof.

This estimate easily implies estimates for thicker slabs. If $\delta < w < 1$, then a slab of the form $N_w(\pi) \cap B(2)$ can be covered by $\sim (w/\delta)^2$ slabs of the form $N_\delta(\pi) \cap B(2)$. Therefore, for any w in the range $\delta < w < 1$, we get the estimate

$$|N_w(\pi) \cap N_\delta(H) \cap B(2)| \lesssim w^2 \delta \log(\delta^{-1}). \tag{15.16}$$

From the estimate 15.16, we see that $|N_w(\pi) \cap N_\delta(H) \cap B(2)|$ can be covered by $\sim w^2 \delta^{-5} \log(\delta^{-1})$ δ-cubes. Each of these cubes lies in $\lesssim \delta^{-1}$ tubes $T_i \subset N_w(\pi)$,

and each tube T_i contains δ^{-2} cubes. Therefore, the number of tubes $T_i \subset N_w(\pi)$ is $\lesssim w^2\delta^{-4}\log(\delta^{-1})$. Up to the log factor, this is the bound in Condition 15.14. It just remains to check the volume estimate in Equation 15.15.

To prove this estimate, we prove general bounds about the volume of the set where a degree 2 polynomial is small. Here is the general question we will consider. Suppose that P is a degree 2 polynomial on $\mathbb{R}^n$. We let $|P|$ denote the maximum of the norms of the coefficients of P. If we normalize so that $|P| \sim 1$, what is the maximum possible volume of $\{x \in B^n(1)$ so that $|P(x)| \leq \delta\}$? For many polynomials P, this volume is $\lesssim \delta$, but there is an exception. If $P(x) = x_1^2$, then the condition $|P(x)| \leq \delta$ is equivalent to $|x_1| \leq \delta^{1/2}$. In this case, the set has volume $\sim \delta^{1/2}$. More generally, if P is the square of a linear polynomial, we get a volume of $\sim \delta^{1/2}$. This is the only way to get a volume near $\delta^{1/2}$.

PROPOSITION 15.43. Suppose that $P \in \mathrm{Poly}_2(\mathbb{R}^n)$ has $|P| \sim 1$ and suppose that for any $P_1 \in \mathrm{Poly}_1(\mathbb{R}^n)$, $|P - P_1^2| \gtrsim 1$. Then for any $0 < \delta < 1/2$,

$$|\{x \in B^n(1) \text{ so that } |P(x)| < \delta\}| \lesssim \delta(\log \delta^{-1}).$$

The expression $\delta(\log \delta^{-1})$ is sharp in this inequality, as we can see from the example $P(x_1, x_2) = x_1^2 - x_2^2$. For any $1 \leq j \leq (\log \delta^{-1})$, the rectangle defined by $|x_1| \sim 2^{-j}$, $|x_2| \sim 2^j\delta$ lies in the set where $|P(x)| \leq \delta$. There are $\sim \log \delta^{-1}$ such rectangles, they are disjoint, and each one has area $\sim \delta$.

PROOF. To prove the proposition, we build up to it from simple cases. The first simple case is the polynomial in one variable $P(x) = \lambda x^2 + b$. If $|\lambda| \sim 1$, then we claim that

$$|\{x \in [-1, 1] \text{ so that } |P(x)| \leq \delta\}| \lesssim \min(|b|^{-1/2}\delta, \delta^{1/2}).$$

If $|b| \leq 2\delta$, then we have the condition $|\lambda x^2| \leq 3\delta$, and so $|x| \lesssim \delta^{1/2}$. If $|b| > 2\delta$, then we proceed using the fundamental theorem of calculus. If $|P(x)|$ is small, then x must be close to $\pm\lambda^{-1/2}|b|^{1/2}$. In this region $|P'(x)| = |2\lambda x| \sim |b|^{1/2}$. Therefore, $P(x)$ changes by δ over an interval of length $\sim |b|^{-1/2}\delta$.

If we know that $|b| \sim 1$, then the left-hand side is $\lesssim \delta$. On the other hand, if we don't know anything about b, then the left-hand side is $\lesssim \delta^{1/2}$.

The second simple case is a polynomial in two variables $P(x_1, x_2) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + b$, with $|\lambda_1|, |\lambda_2| \sim 1$ and $|b| \lesssim 1$. In this case, we claim that

$$\big|\{(x_1, x_2) \in B^2(1) \text{ so that } |P(x)| \leq \delta\}\big| \lesssim \delta \log(\delta^{-1}).$$

We organize the set on the left according to the order of magnitude of $\lambda_2 x_2^2 + b$. For $1 \leq j \leq \log \delta^{-1}$, we define

$$S_j := \{x_2 \in [-1, 1] \text{ so that } |\lambda_2 x_2^2 + b| \leq 2^{-j}.$$

Since $|\lambda_2| \sim 1$, the first simple case tells us that $|S_j| \lesssim 2^{-j/2}$.

Suppose $x_2 \in S_j$ but $x_2 \notin S_{j+1}$, we have $|\lambda_2 x_2^2 + b| \sim 2^{-j}$. Since $|\lambda_1| \sim 1$, the first simple case tells us that

$$|\{x_1 \in [-1, 1] \text{ so that } |P(x_1, x_2)| \leq \delta\}| \lesssim 2^{j/2}\delta.$$

Therefore, the area we want to bound is at most

$$\sum_{j=1}^{\log \delta^{-1}} 2^{j/2}\delta |S_j|.$$

Plugging in our estimate $|S_j| \lesssim 2^{-j/2}$, we get the desired bound

$$\big|\{(x_1, x_2) \in B^2(1) \text{ so that } |P(x)| \leq \delta\}\big| \lesssim \delta \log(\delta^{-1}).$$

Now we are ready to tackle the general case. After making a rotation of the coordinates, we can assume that the second-order part of P is diagonal:

$$P(x) = \sum_{i=1}^{n} \lambda_i x_i^2 + \sum_{i=1}^{n} b_i x_i + c.$$

The norm $|P|$ is not exactly invariant under rotation of the coordinates, but it is straightforward to check that it is approximately invariant, so we still have $|P| \sim 1$.

We order the coordinates so that $|\lambda_1| \geq |\lambda_2| \geq ...$ If two of the λ's are ~ 1, then we can reduce the problem to our second simple case. For any fixed $(x_3, ..., x_n) \in B(1)$, we can bound the volume of $\{(x_1, x_2) \in B(1) \text{ so that } |P(x)| \leq \delta\}$ by the second simple case, because after translating (x_1, x_2) by a vector of norm $\lesssim 1$, this set is equivalent to a set of the form

$$\{(y_1, y_2) \in B(C) \text{ so that } |\lambda_1 y_1^2 + \lambda_2 y_2^2 + b'| \leq \delta\}.$$

Next suppose that $|\lambda_1| \sim 1$, but the other λ_i are all close to zero. By translating in the x_1-direction, we can assume that $b_1 = 0$.

If $|b_j| \sim 1$ for some $j \geq 2$, then we proceed as follows. Without loss of generality, we can assume that $b_j > 0$. Since λ_j is almost zero, $\partial_j P(x) \sim 1$ for $x \in [-1, 1]^n$. Therefore, for any values of $x_1, ..., x_{j-1}, x_{j+1}, ..., x_n$ in $[-1, 1]$, we get the bound

$$|\{x_j \in [-1, 1] \text{ so that } |P(x)| \leq \delta\}| \lesssim \delta.$$

This gives the desired bound in this scenario.

We continue to suppose that $|\lambda_1| \sim 1$, that the other λ_i are close to zero, and that $b_1 = 0$. If $|b_j|$ is close to zero for all j, then we must have $|c| \sim 1$, because $P(x)$ cannot be too close to the polynomial $\lambda_1 x_1^2$, which is a square. Now, for each $(x_2, ..., x_n) \in B^{n-1}(1)$ the equation $|P(x)| \leq \delta$ expands out to $|\lambda_1 x_1^2 + c'| \leq \delta$ for $|c'| \sim 1$. The set of x_1 satisfying the inequality has length $\lesssim \delta$ which again gives a good bound. This finishes the case where exactly one λ_i has norm ~ 1.

Finally, the case where all λ_i are close to zero is fairly easy. If all $|\lambda_i|$ and all $|b_i|$ are close to zero, then we must have $|c| \sim 1$, and the set of solutions $\{x \in B^n(1) \text{ so that } |P(x)| \leq \delta\}$ is just empty.

Suppose that for some i, λ_i is close to zero, but b_i is far from zero. Without loss of generality, suppose that $i = 1$ and suppose that $b_1 > 0$, and so $b_1 \sim 1$. Then $\partial_1 P(x) \sim 1$ for all $x \in [-1, 1]^n$. Therefore, for any $(x_2, ..., x_n) \in B(1)$,

$$|\{x_1 \in [-1, 1] \text{ so that } |P(x)| \leq \delta\}| \lesssim \delta.$$

This gives the required bound in this case.

□

Now we apply this Proposition to understand the size of $N_\delta(\pi) \cap N_\delta(H) \cap B(2)$.

Suppose that the plane π is given by the equation

$$z_3 = a_1 z_1 + a_2 z_2 + b,$$

where $a_1, a_2, b \in \mathbb{C}$, and also suppose that $|a_1|, |a_2|, |b| \lesssim 1$. An arbitrary plane π that passes through $B(2)$ can be written in this form after possible exchanging the coordinates. The argument that we will make applies equally well if the variable on the left-hand side is z_1 or z_2.

If $z \in N_\delta(\pi)$, then we have

$$|z_3 - a_1 z_1 - a_2 z_2 - b| \leq C\delta. \tag{15.17}$$

Also, if $z \in N_\delta(H) \cap B(2)$, we have

$$\left||z_3|^2 - |z_1|^2 - |z_2|^2 + 1\right| \leq C\delta. \tag{15.18}$$

This inequality requires a little explanation. Let $F(z_1, z_2, z_3) = |z_1|^2 + |z_2|^2 - |z_3|^2$, so that H is just $\{z \in \mathbb{C}^3 \text{ so that } F(z) = 1\}$. We view F as a function from $\mathbb{R}^6$ to $\mathbb{R}$. The gradient of F has a simple formula, from which we see that $\nabla F(z) = 0$ only at $z = 0$. Here is the formula:

$$\nabla F(z) = (\partial_{x_1} F, \partial_{y_1} F, \partial_{x_2} F, \partial_{y_2} F, \partial_{x_3} F, \partial_{y_3} F) = (2x_1, 2y_1, 2x_2, 2y_2, -2x_3, -2y_3).$$

Since $0 \notin H$, we have $|\nabla F(z)| \sim 1$ for $z \in H \cap B(2)$. Therefore, for all $z \in B(2)$, $|F(z)| \sim \mathrm{Dist}(z, H)$. So for all $z \in N_\delta(H) \cap B(2)$, we have $\left||z_3|^2 - |z_1|^2 - |z_2|^2 - 1\right| = |F(z)| \leq C\delta$.

We now combine Equations 15.17 and 15.18. Suppose $(z_1, z_2, z_3) \in N_\delta(\pi) \cap N_\delta(H) \cap B(2)$. By Equation 15.17

$$|z_3|^2 = |a_1 z_1 + a_2 z_2 + b|^2 + O(\delta).$$

But by Equation 15.18, we have

$$|z_3|^2 = |z_1|^2 + |z_2|^2 - 1 + O(\delta).$$

Hence we get an inequality about z_1, z_2:

$$|a_1 z_1 + a_2 z_2 + b|^2 = |z_1|^2 + |z_2|^2 - 1 + O(\delta). \tag{15.19}$$

We let X denote the set of $(z_1, z_2) \in B(2)$ obeying this inequality. We will prove that $|X| \lesssim \delta \log(\delta^{-1})$. This bound implies Equation 15.15: if $(z_1, z_2, z_3) \in N_\delta(\pi) \cap N_\delta(H) \cap B(2)$, then $(z_1, z_2) \in X$, and for any $(z_1, z_2) \in X$, the set of $z_3 \in \mathbb{C}$ so that $(z_1, z_2, z_3) \in N_\delta(\pi)$ lies is a disk of radius $\lesssim \delta$.

To see that $|X| \lesssim \delta \log(\delta^{-1})$, we will apply Proposition 15.43. The set X is defined by the inequality

$$|P(x_1, y_1, x_2, y_2)| \leq C\delta,$$

where

$$P(x_1, y_1, x_2, y_2) = |a_1 z_1 + a_2 z_2 + b|^2 - |z_1|^2 - |z_2|^2 + 1.$$

Expanding out z_i in terms of x_i and y_i, we see that P is a degree 2 polynomial with $|P| \lesssim 1$. To apply Proposition 15.43, we just have to check that for any degree 1 polynomial P_1, $|P - P_1^2| \gtrsim 1$. This is not hard to check by expanding out P a little bit. We focus on the degree 2 terms of P. The degree 2 part of P is

$$(|a_1|^2 - 1)|z_1|^2 + (|a_2|^2 - 1)|z_2|^2 + 2\Re(a_1 \bar{a}_2 z_1 \bar{z}_2).$$

When we expand in terms of x_i, y_i, we see that

$$|z_i|^2 = x_i^2 + y_i^2,$$

and

$$z_1 \bar{z}_2 = (x_1 x_2 + y_1 y_2) + i(-x_1 y_2 + x_2 y_1).$$

Therefore, the real part of $a_1\bar{a}_2 z_1 \bar{z}_2$ has the form

$$\Re(a_1\bar{a}_2 z_1 \bar{z}_2) = b_1(x_1x_2 + y_1y_2) + b_2(-x_1y_2 + x_2y_1).$$

If $|a_1|, |a_2| \sim 1$, then $(b_1, b_2) \sim 1$ also.

Now we can check that P is not too close to a square. If $(|a_1|^2 - 1) \gtrsim 1$, then P has large x_1^2 and y_1^2 coefficients but thex_1y_1 coefficient of P is zero, and so $|P - P_1^2| \gtrsim 1$. Similarly if $(|a_2|^2 - 1)$ is not too small. So we are left with the case that $|a_1|$ and $|a_2|$ are very close to 1. In this case, the x_i^2 and y_i^2 coefficients of P are all small, but some of the other coefficients have size ~ 1. This also guarantees that $|P - P_1^2| \gtrsim 1$. If P_1 has a significant coefficient in front of one of the degree 1 terms, say cx_2, then the x_2^2-coefficient of P_1^2 will be ~ 1, much different from the x_2^2-coefficient of P. But if all the degree 1 coefficients of P_1 are small, then all the degree 2 coefficients of P_1^2 will also be small, and so $|P - P_1^2| \gtrsim 1$ in that case as well.

Therefore, P obeys the hypotheses of Proposition 15.43, and we get the desired bound $|X| \lesssim \delta \log(\delta^{-1})$. This finishes the proof of our slightly weaker version of Theorem 15.40.

Exercise 15.13. The Hermitian variety has been studied a lot in complex differential geometry. One interesting property is that it contains no 2-dimensional complex submanifolds - not even very small ones. Every complex manifold is locally the graph of a holomorphic function. So consider the graph of a holomorphic function f,

$$z_3 = f(z_1, z_2),$$

where f is defined for (z_1, z_2) in an open set $\Omega \subset \mathbb{C}^2$. If the graph of f lies in H, then we get the equation

$$|z_1|^2 + |z_2|^2 - |f(z_1, z_2)|^2 = 1.$$

Show that no holomorphic function obeys this equation.

Hint: differentiate the equation and use the fact that f is holomorphic. It may be helpful to recall the definition of ∂_j and $\bar{\partial}_j$ in complex analysis:

$$\partial_j := \frac{1}{2}(\partial_{x_j} - i\partial_{y_j}),$$

$$\bar{\partial}_j = \frac{1}{2}(\partial_{x_j} + i\partial_{y_j}).$$

For example, $\partial_j z_j = 1$, and $\bar{\partial}_j z_j = 0$. Similarly, $\partial_j \bar{z}_j = 0$ and $\bar{\partial}_j \bar{z}_j = 1$. A function f is holomorphic if and only if $\bar{\partial}_j f = 0$ for all j. The operators ∂_j and $\bar{\partial}_j$ obey the Leibniz formula. They also behave nicely with respect to complex conjugation: for any function f,

$$\overline{\partial_j f} = \bar{\partial}_j \bar{f}.$$

https://doi.org/10.1090//ulect/064/16

CHAPTER 16

The polynomial method in number theory

In the early 20th century, Thue made an important breakthrough in the study of diophantine equations. The arguments we have explored in this book have a lot of similarities to his argument from 1909. In this chapter, we will prove Thue's theorem about diophantine equations and note the parallels to other arguments in the book. Here is the statement of the theorem.

THEOREM 16.1. (Thue) Suppose $P \in \mathbb{Z}[x, y]$ is a homogeneous polynomial with degree ≥ 3 which is irreducible over $\mathbb{Z}$. If A is any integer, then the equation $P(x, y) = A$ has only finitely many integer solutions.

This theorem is important because it applies to a much more general class of diophantine equations than any previous theorem. Also the method of proof influenced a lot of later work in number theory, including diophantine equations, transcendental number theory, and some work on exponential sums. There is a third reason that I think the theorem is important, which I will explain in the next section.

The proof of this theorem is the main topic of the chapter, but before we turn to the proof we take some time to put the result in context.

16.1. Naive guesses about diophantine equations

Suppose that P is a polynomial of degree d in n variables, with integer coefficients. Let us consider the equation $P(x) = 0$. We would like to make an educated guess about the number of solutions of this equation. The total number of solutions may be infinite, and in this case, it's interesting to try to estimate the number of solutions in a given size range. To be precise, let us try to estimate the size of the set

$$\{x \in \mathbb{Z}^n \text{ so that } P(x) = 0 \text{ and } 2^s \leq |x| < 2^{s+1}\}.$$

We note that if $|x| \sim 2^s$, then $|P(x)| \lesssim 2^{sd}$. Just based on this simple observation, we can make a primitive probabilistic model: For each x with $2^s \leq |x| < 2^{s+1}$, let $\tilde{P}(x)$ be a random integer in the range $[-2^{ds}, 2^{ds}]$. The number of $x \in \mathbb{Z}^n$ with $2^s \leq |x| < 2^{s+1}$ is $\sim 2^{ns}$. Therefore, the expected size of the set $\{x \in \mathbb{Z}^n \text{ so that } \tilde{P}(x) = 0 \text{ and } \}$ is $\sim 2^{ns}/2^{ds} = 2^{(n-d)s}$. If the polynomial P "behaved randomly", then the number of solutions to $P(x) = 0$ with $|x| \sim 2^s$ would be $\sim 2^{(n-d)s}$. This suggests the following naive conjectures.

NAIVE CONJECTURE 1. Suppose that P is a polynomial in n variables with integer coefficients. If $\operatorname{Deg} P = d \leq n$, then the equation $P(x) = 0$ has infinitely many integer solutions, and the number of solutions of size $\sim 2^s$ is $\sim 2^{(n-d)s}$.

NAIVE CONJECTURE 2. Suppose that P is a polynomial in n variables with integer coefficients. If $\operatorname{Deg} P > n$, then the equation $P(x) = 0$ has only finitely many integer solutions.

These conjectures are both false, but they still offer a useful perspective. Let us give some counterexamples using polynomials in two variables.

Consider the equation $2x + 2y - 1 = 0$. Our model predicts it should have many solutions, but it has none because the left-hand side is always even but the right-hand side is always odd. Therefore, Naive Conjecture 1 is false. Now we turn to naive conjecture 2, which is more closely related to Thue's work.

We consider the equation $(x - y)^9 - 1 = 0$. It has degree 10, so our model predicts that it should have only finitely many solutions, but it has infinitely many solutions, because every solution of $x - y = 1$ is a solution. This shows that Naive Conjecture 2 is false. The polynomial $(x - y)^9 - 1$ is reducible over the complex numbers, and one of the factors is $x - y - 1$, which is a polynomial of degree 1. But Naive Conjecture 2 is false also for irreducible polynomials.

Here is a subtler counterexample to Naive Conjecture 2. Consider a polynomial map from $\mathbb{R}$ to $\mathbb{R}^2$ with integer coefficients, for example:

$$\phi(t) = (t^2 + 1, t^3 + t + 1).$$

As we saw in Exercise 6.2, the image of such a polynomial map lies in the zero set of a polynomial of two variables: $P(x, y) = 0$. Moreover, we can arrange that P has integer coefficients. In this case, the equation $P(x, y) = 0$ will have infinitely many integer solutions, because for every integer t, $\phi(t)$ will be an integer solution of the equation $P(x, y) = 0$. The degree of P can be made arbitrarily large by choosing $\phi(t)$ in a complicated way. Also, if we choose the polynomial P with lowest possible degree vanishing on the image of ϕ, then P will be irreducible.

In spite of these counterexamples, the naive conjectures above are true in many cases. In particular, Thue's theorem says that Naive Conjecture 2 is true for polynomials in two variables of the form $P(x, y) = A$ if A is an integer and P is a homogeneous polynomial which is irreducible over $\mathbb{Z}$. As far as I know, Thue's theorem was the first major result confirming this probabilistic intuition. This is the third reason I think Thue's theorem is important.

To end this section, we briefly discuss the condition that P is irreducible over $\mathbb{Z}$. Recall that a polynomial $P \in \mathbb{Z}[x, y]$ is reducible over $\mathbb{Z}$ if $P = P_1 \cdot P_2$ where $P_1, P_2 \in \mathbb{Z}[x, y]$ and each P_i has degree at least 1. If $P(x, y)$ is reducible over $\mathbb{Z}$, then the equation $P(x, y) = A$ actually becomes much easier to understand. We give a sense of the reducible case by describing two examples.

Our first example is the equation

$$(x^3 - 2y^3)(x - 3y) = 12. \tag{16.1}$$

If $(x, y) \in \mathbb{Z}^2$ solve this equation, then there must be *integers* a and b with $ab = 12$ so that

$$x^3 - 2y^3 = a, x - 3y = b. \tag{16.2}$$

There are only finitely many choices for (a, b). For each choice of (a, b), there are only finitely many complex solutions (x, y) to Equation 16.2. So we see that there are only finitely many integer solutions to Equation 16.1. And we could also use this argument to systematically find them all.

In our second example, $P(x, y)$ is a power of another polynomial. Suppose that $P_1(x, y)$ is homogeneous and irreducible over $\mathbb{Z}$ and consider the equation

$$P_1(x, y)^{10} = 2^{10}. \tag{16.3}$$

This equation is equivalent to the equation $P_1(x, y) = \pm 2$. So this second example reduces to a simpler equation. If the degree of P_1 is at least 3, then Thue's theorem implies that there are only finitely many solutions. (In the next section, we will study an example where the degree of P_1 is 2.)

16.2. Parabolas, hyperbolas, and high degree curves

To keep building our intuition, we now study the integer solutions to some simple diophantine equations. We start by comparing the parabola

$$y - 2x^2 = 1,$$

with the hyperbola

$$y^2 - 2x^2 = 1.$$

We will see that both the parabola and the hyperbola have infinitely many integer points, but we will also see that they are distributed very differently: the integer points on the hyperbola are much sparser than the integer points on the parabola. For any integer x, the point $(x, 2x^2 + 1)$ lies in the parabola. Therefore, the parabola contains $\sim N^{1/2}$ integer points (x, y) with $|x|, |y| \leq N$. This is far more integer points than predicted by the naive conjectures in the last section. On the other hand, we will show that the hyperbola obeys the naive conjectures from the last section.

For any integer $s \geq 0$, define

$$\mathrm{Sol}_s := \{(x, y) \in \mathbb{Z}^2 \text{ so that } y^2 - 2x^2 = 1 \text{ and } 20^s \leq |x| \leq 20^{s+1}\}.$$

The set Sol_s is the set of integer points on the hyperbola at scale 20^s. The number of points in Sol_s is controlled by the following proposition:

PROPOSITION 16.2. For every integer $s \geq 0$,

$$|\,\mathrm{Sol}_s\,| \sim 1.$$

(We used 20^s instead of 2^s in the definition of Sol_s so that we will be able to prove that Sol_s contains at least one point for every $s \geq 0$.)

We are most interested in the upper bound in this Proposition, which shows that the hyperbola has far fewer integer points than the parabola. The key observation is that for any integer point on the hyperbola, y/x is very close to $\sqrt{2}$ or to $-\sqrt{2}$. We will focus on solutions where $x, y > 0$, so that y/x is very close to $\sqrt{2}$. We make this precise in the following lemma:

LEMMA 16.3. If x and y are positive integers with $y^2 - 2x^2 = 1$, then

$$\left|\frac{y}{x} - \sqrt{2}\right| \leq |x|^{-2}.$$

PROOF. Starting with $y^2 - 2x^2 = 1$ and dividing by x^2, we get

$$\left|(\frac{y}{x})^2 - 2\right| \leq |x|^{-2}.$$

Now we factor the left-hand side to get

$$\left|\frac{y}{x} - \sqrt{2}\right| \cdot \left|\frac{y}{x} + \sqrt{2}\right| \leq |x|^{-2}.$$

Since $x, y > 0$ we have $|\frac{y}{x} + \sqrt{2}| > 1$. □

If $(x, y) \in \mathrm{Sol}_s$, then the ratio y/x lies in an interval around $\sqrt{2}$ of length $\sim 20^{-2s}$. Next, we will show that all these ratios are different.

LEMMA 16.4. If $(x, y) \in \mathbb{Z}^2$ and $y^2 - 2x^2 = 1$, then $\gcd(x, y) = 1$.

PROOF. If a divides both x and y, then a divides $y^2 - 2x^2 = 1$. □

Finally, if (x_1, y_1) and (x_2, y_2) are in Sol_s, then the ratios y_1/x_1 and y_2/x_2 cannot be too close together.

LEMMA 16.5. If $(x_1, y_1), (x_2, y_2) \in \mathrm{Sol}_s$, then

$$\left|\frac{y_1}{x_1} - \frac{y_2}{x_2}\right| \gtrsim 20^{-2s}.$$

PROOF. By the previous lemma, y_1/x_1 and y_2/x_2 are in lowest terms, and so the left-hand side is non-zero. We write the left-hand side as

$$\left|\frac{y_1 x_2 - y_2 x_1}{x_1 x_2}\right|.$$

Since the left-hand side is non-zero, it must be at least $\frac{1}{x_1 x_2}$. Since $|x_1|, |x_2| \leq 20^{s+1}$,

$$\frac{1}{x_1 x_2} \geq [20^{s+1}]^{-2} \gtrsim 20^{-2s}.$$

□

We can now prove the upper bound on $|\mathrm{Sol}_s|$. We have $|\mathrm{Sol}_s|$ different points $(x, y) \in \mathrm{Sol}_s$, and so we get $|\mathrm{Sol}_s|$ different ratios y/x. These ratios are all contained in an interval around $\sqrt{2}$ of length $\sim 20^{-2s}$, and the distance between any two ratios is at least $\sim 20^{-2s}$. Therefore, $|\mathrm{Sol}_s| \lesssim 1$.

Geometrically, the difference between the parabola and the hyperbola is that the hyperbola has asymptotic lines, $y = \pm\sqrt{2}x$. Because of these asymptotic lines, for any point (x, y) on the hyperbola with $|x|$ large, y/x is close to $\pm\sqrt{2}$. On the other hand, the parabola does not have asymptotic lines, and so we don't get the same amount of control over the ratio y/x for points in the parabola.

Finding integer points on the hyperbola $y^2 - 2x^2 = 1$ is also based on the observation that y/x should be close to $\sqrt{2}$. We consider the continued fraction expansion of $\sqrt{2}$. The first few terms are $1, \frac{3}{2}, \frac{7}{5}, \frac{17}{12}, \ldots$ Every other term in this continued fraction expansion corresponds to an integer solution of $y^2 - 2x^2 = 1$. The second term gives us $3^2 - 2 \cdot 2^2 = 1$, and the fourth term gives us $17^2 - 2 \cdot 12^2 = 289 - 2 \cdot 144 = 1$. The other terms give solutions to $y^2 - 2x^2 = -1$. The continued fraction expansion can be defined recursively by setting $x_1 = y_1 = 1$, $x_2 = 3, y_2 = 2$, and then defining

$$x_{n+1} = 2x_n + x_{n-1}; y_{n+1} = 2y_n + y_{n-1}.$$

EXERCISE 16.1. With x_n and y_n defined as above, show that

$$y_n^2 - 2x_n^2 = (-1)^n.$$

This exercise finishes the proof of Proposition 16.2.

The argument for the hyperbola also gives us a certain amount of insight into higher-degree curves. For instance, let us consider the curve

$$y^d - 2x^d = 1,$$

for some degree $d \geq 3$. For positive integer solutions to this equation y/x is close to $2^{1/d}$, and the distance $|\frac{y}{x} - 2^{1/d}|$ is much smaller than for a hyperbola. We state this estimate as a lemma.

LEMMA 16.6. If $x, y > 0$ are integers and $y^d - 2x^d = 1$, then

$$\left|\frac{y}{x} - 2^{1/d}\right| \lesssim |x|^{-d}.$$

The proof of this Lemma is essentially the same as the proof of Lemma 16.3, so we omit the details. Now let $(x_1, y_1), (x_2, y_2), ...$ be all the pairs of positive integers obeying $y_d - 2x^d = 1$, ordered so that $x_1 < x_2 < ....$ Lemma 16.6 leads to some interesting estimates about the behavior of x_j.

PROPOSITION 16.7. For $d \geq 3$, we have

$$x_j \gtrsim x_{j-1}^{d-1}.$$

Therefore, the number of positive integer solutions (x, y) with $x < N$ is $\lesssim \log\log N$.

Recall that the equation $y^2 - 2x^2 = 1$ has $\sim \log N$ positive integer solutions (x, y) with $x < N$. So Proposition 16.7 shows that the curve $y^3 - 2x^3 = 1$ has far fewer integer points than the curve $y^2 - 2x^2 = 1$. (For simplicity, we discussed positive integer points in the statement of Proposition 16.7, but similar arguments apply if x or y is negative.)

PROOF OF PROPOSITION 16.7. If $y^d - 2x^d = 1$, then $\gcd(x, y) = 1$. Therefore, y_j/x_j are all fractions in lowest terms. Since $|\frac{y_j}{x_j} - 2^{1/d}| \lesssim |x_j|^{-d}$, we see that

$$\left|\frac{y_j}{x_j} - \frac{y_{j-1}}{x_{j-1}}\right| \lesssim x_{j-1}^{-d}.$$

On the other hand, the left-hand side is a non-zero fraction with denominator $x_{j-1}x_j$, and so

$$\frac{1}{x_{j-1}x_j} \lesssim x_{j-1}^{-d}.$$

Rearranging, we get $x_j \gtrsim x_{j-1}^{d-1}$. This gives the desired bound on x_j. Next, we would like to iterate this inequality to estimate x_j. There must be some j_0 so that for all $j \geq j_0$, we have $x_j \geq x_{j-1}^{3/2}$. Also $x_{j_0} \geq 2$. Therefore,

$$x_j \geq 2^{(\frac{3}{2})^{j-j_0}}.$$

If $x_j \leq N$, then we see that $j - j_0 \lesssim \log\log N$. Also j_0 is independent of N, and so we get $j \lesssim \log\log N$. □

Proposition 16.7 shows that for $d \geq 3$, the curve $y^d - 2x^d = 1$ contains far fewer integer points than the hyperbola $y^2 - 2x^2 = 1$. It shows that the number of integer points on $y^d - 2x^d = 1$ with $|(x, y)| \leq N$ is $\lesssim \log\log N$. But this bound is not sharp. Thue's theorem shows that the number of integer points on the curve is finite, consistent with our probabilistic intuition from the last section.

Thue's theorem is subtler than the argument we have seen so far. In particular, it takes a new idea to rule out the possibility that there are infinitely many solutions $(x_1, y_1), (x_2, y_2), ...$ at wildly different scales. We will see how Thue did this using a version of the polynomial method.

EXERCISE 16.2. Suppose that $P(x, y)$ is a polynomial with integer coefficients with degree $d \geq 2$. Suppose that P has no terms of degree $d-1$, so what we can write $P(x, y) = P_d(x, y) + P_{\leq d-2}(x, y)$, where P_d is homogeneous of degree d and $P_{\leq d-2}$ has degree at most $d-2$. Suppose that P_d is irreducible over $\mathbb{Z}$. Prove that the zero set of P contains $\lesssim \log N$ integer points of size $\leq |N|$.

16.3. Diophantine approximation

We saw in the last section that for any integer solution to $y^d - 2x^d = 1$, the ratio y/x must be a very good rational approximation to $2^{1/d}$. In order to prove Theorem 16.1, Thue proved a general theorem about rational approximations of algebraic numbers. His theorem states that there are only finitely many "very good" rational approximations to an algebraic number.

Before stating Thue's theorem, we discuss some simpler results about diophantine approximation in order to put it in perspective. The question we will investigate is the following: given a real number β and a parameter $s > 0$, how many rational numbers $\frac{p}{q}$ obey the inequality

$$\left|\beta - \frac{p}{q}\right| \leq |q|^{-s}.$$

When $s = 2$, there are infinitely many solutions to this inequality for any irrational number β.

PROPOSITION 16.8. (Dirichlet) For any irrational real number β, there are infinitely many solutions to the diophantine inequality

$$\left|\beta - \frac{p}{q}\right| \leq |q|^{-2}.$$

PROOF. For any real number x, let $\langle x \rangle \in [0, 1)$ be the fractional part of x. Let Q be any number, and consider the sequence of numbers $\langle \beta \rangle, \langle 2\beta \rangle, ..., \langle Q\beta \rangle$ in $[0, 1)$. Since there are Q of these numbers, it must be possible to find two of them within a distance $1/Q$. In other words, we can choose $1 \leq q_1 < q_2 \leq Q$ so that the distance from $\langle q_1\beta \rangle$ to $\langle q_2\beta \rangle$ is at most $1/Q$. This implies that

$$|\langle (q_2 - q_1)\beta \rangle| \leq 1/Q.$$

We let $q = q_2 - q_1$. The last inequality implies that there is some integer p so that

$$|q\beta - p| \leq 1/Q.$$

Dividing by $q \leq Q$, we see that

$$\left|\beta - \frac{p}{q}\right| \leq \frac{1}{qQ} \leq q^{-2}.$$

By increasing Q, we get infinitely many solutions to the diophantine inequality $\left|\beta - \frac{p}{q}\right| \leq |q|^{-2}$. □

On the other hand, for $s > 2$, almost every real number β admits only a finite number of solutions to the inequality $\left|\beta - \frac{p}{q}\right| \leq |q|^{-s}$. This result is a standard exercise in measure theory.

EXERCISE 16.3. Let D_s be the set of numbers $\beta \in \mathbb{R}$ so that there are infinitely many rational solutions to the diophantine inequality

$$\left|\beta - \frac{p}{q}\right| \leq |q|^{-s}.$$

Prove that for any $s > 2$, the set D_s has measure zero.

Dirichlet's result and this exercise together give a good understanding of what happens for a "typical" real number. But it can be much harder to understand what happens for any particular number, such as π or $2^{1/3}$. In this book, we won't discuss π, but the example of $\beta = 2^{1/3}$ will be very important for us. The first estimate about diophantine approximation of algebraic numbers was given by Liouville in the 1840's.

PROPOSITION 16.9. (Liouville) If β is an irrational algebraic number of degree $d \geq 2$ and $\frac{p}{q}$ is a rational number, then

$$|\beta - \frac{p}{q}| \geq c(\beta)|q|^{-d}.$$

PROOF. Suppose that Q is a minimal degree polynomial with integer coefficients obeying $Q(\beta) = 0$. By Taylor's theorem, for all z in the interval $|\beta - z| \leq 1$, we have an inequality of the form

$$|Q(z)| \leq C(\beta)|\beta - z|,$$

for some constant $C(\beta)$. (The value of $C(\beta)$ depends on the size of $|Q'|$ and $|Q''|$ on the interval $|\beta - z| \leq 1$.)

Now we consider a rational number p/q with $|\beta - (p/q)| \leq 1$. By the last indented equation, we know that

$$|Q(p/q)| \leq C(\beta)\left|\beta - \frac{p}{q}\right|.$$

On the other hand, since Q is a polynomial with integer coefficients of degree d, we know that $Q(p/q)$ is a rational number with denominator q^d. Therefore, either $Q(p/q) = 0$ or else $|Q(p/q)| \geq |q|^{-d}$. Since Q is the minimal degree integer polynomial vanishing at β, it turns out that $Q(p/q)$ cannot be zero. This follows from Gauss's Lemma, and we indicate the proof in Exercise 16.7 below. Therefore $|Q(p/q)| \geq |q|^{-d}$, and so we get

$$|q|^{-d} \leq C(\beta)\left|\beta - \frac{p}{q}\right|.$$

This finishes the proof of the Proposition. (We remark that the constant $c(\beta)$ works out to $C(\beta)^{-1}$, which depends on the size of $|Q'|$ and $|Q''|$ near β.)

□

For $d = 2$, Liouville's inequality and Dirichlet's proposition match each other closely: there are infinitely many solutions to $|\beta - (p/q)| \leq |q|^{-2}$ but no solutions to $|\beta - (p/q)| \leq c(\beta)|q|^{-2}$. But for $d = 3$, Liouville's inequality and Dirichlet's proposition do not match each other so closely. There are infinitely many solutions

to $|\beta - (p/q)| \leq |q|^{-s}$ when $s = 2$, and Liouville implies that there are only finitely many solutions when $s > 3$. But what about $2 < s \leq 3$? Recall that for almost every real number β, there are only finitely many solutions to $|\beta - (p/q)| \leq |q|^{-s}$ for any $s > 2$. But what happens for $\beta = 2^{1/3}$?

(It might be worth remarking here that there is a good way to gather experimental numerical evidence about this question. The method of continued fractions gives an efficient algorithm to compute the best rational approximations p/q of a number β with $q \leq Q$. Before computers, it was already practical to find the best rational approximations to $2^{1/3}$ with denominator up to, say, 10^{20}.)

In the early 1900's, Thue proved a theorem on diophantine approximation that strengthens Liouville's inequality. As a corollary, this theorem implies Theorem 16.1 on diophantine equations.

THEOREM 16.10. Suppose that β is an algebraic number of degree $d \geq 3$, and suppose that $s > \frac{d+2}{2}$. Then there are only finitely many rational numbers p/q that satisfy the inequality

$$\left|\beta - \frac{p}{q}\right| \leq q^{-s}.$$

For instance, if $d = 3$ and $s > 2.5$, then this theorem implies that there are only finitely many solutions to the inequality $|\beta - (p/q)| \leq q^{-s}$. The range of s in Theorem 16.10 is not sharp, but the important point is that it improves on Liouville's estimate. As we will see even a tiny improvement on Liouville's estimate is enough to prove Theorem 16.1, and so even a tiny improvement is very interesting.

Let us see how Theorem 16.10 implies that some diophantine equations have only finitely many solutions. We begin with the equation $y^d - 2x^d = 1$, for $d \geq 3$, because the proof is particularly simple. As we observed in the last section, any solution to this equation has $\gcd(x, y) = 1$ and obeys the inequality $|2^{1/d} - (y/x)| \leq C_d x^{-d}$. So integer solutions correspond to rational numbers that are very good approximations of $2^{1/d}$. Theorem 16.10 implies that there exists an exponent $s < d$ so that there are only finitely many rational numbers obey $|2^{1/d} - (y/x)| \leq |x|^{-s}$. For instance, we can take $s = d - (1/10)$. Now suppose that the equation $y^d - 2x^d = 1$ has infinitely many solutions. With finitely many exceptions, they must all obey:

$$|x|^{-s} \leq |2^{1/d} - (y/x)| \leq C_d |x|^{-d}.$$

Since $s < d$, these inequalities give a bound on $|x|$: $|x| \leq C_d^{\frac{1}{d-s}}$. For any fixed x, it is elementary to see that the equation $y^d - 2x^d = 1$ has only finitely many solutions in y. Therefore, we see that the equation $y^d - 2x^d = 1$ has only finitely many solutions in total.

Essentially the same argument applies in the general setting of Theorem 16.1.

PROOF OF THEOREM 16.1 USING THEOREM 16.10. Suppose that $P(x, y)$ is a homogenous polynomial of degree $d \geq 3$ which is irreducible over $\mathbb{Z}$. We consider solutions to the equation $P(x, y) = A$.

We expand $P(x, y) = \sum_{j=0^d} a_j x^{d-j} y^j$. Dividing on both sides by x^d we see that $\left|\sum_{j=0}^d a_j (y/x)^j\right| = |Ax^{-d}|$. We define $Q(z) = \sum_{j=0}^d a_j z^j$, so that we can write:

$$|Q(y/x)| = |A||x|^{-d}. \tag{16.4}$$

So we see that if (x, y) is an integer solution to $P(x, y) = A$ with $|x|$ large, then $|Q(y/x)|$ is small. We will use this to show that y/x must lie near to one of the roots of Q. Let β_j be the roots of Q.

Since Q is irreducible over the integers, each root appears with multiplicity 1, and so $Q'(\beta_j) \neq 0$. We include a proof of this fact in Exercise 16.8 at the end of the chapter.

Now let us argue more precisely that if (x, y) is an integer solution to $P(x, y) = A$ with $|x|$ large, then y/x must be close to one of the roots of Q. We choose some small number $\delta > 0$ so that for z in the interval of length δ centered at each root β_j, we have the inequality

$$(1/2)|Q'(\beta_j)||z - \beta_j| \leq |Q(z)| \leq 2|Q'(\beta_j)||z - \beta_j|. \tag{16.5}$$

If y/x is not in any of these δ-intervals, then we get a lower bound for $|Q(y/x)|$, and hence an upper bound for $|x|$. So there are only finitely many solutions where y/x is not in any of these δ-intervals.

Fix a root β_j, and consider solutions (x, y) so that y/x lies in the δ-neighborhood centered at β_j. It suffices to show that there are only finitely many such solutions. Combining equation 16.4 and equation 16.5, we see that

$$(1/2)|Q'(\beta_j)| \left|\beta_j - \frac{y}{x}\right| \leq |A||x|^{-d} \leq 2|Q'(\beta_j)||\beta_j - \frac{y}{x}|.$$

In particular, we see that $|\beta_j - (y/x)| \lesssim |x|^{-d}$. Since $d \geq 3$, we can again choose s in the range $\frac{d+2}{2} < s < d$. By Thue's diophantine approximation theorem, Theorem 16.10, we know that there are only finitely many solutions to the inequality $|\beta_j - (y/x)| \leq |x|^{-s}$. So with finitely many exceptions, we get the inequalities

$$|x|^{-s} < \left|\beta_j - \frac{y}{x}\right| \leq C|x|^{-d}.$$

In this equation, the constant C depends on A and on $|Q'(\beta_j)|$, but not on $|x|$. Since $d > s$, these inequalities give us an upper bound on $|x|$. But there are only finitely many solutions with $|x|$ below this bound. □

It is also interesting to consider the best exponent s in Theorem 16.10. As we discussed above, for almost every real number β, and for every $s > 2$, there are only finitely many solutions to the inequality $|\beta - (p/q)| \leq |q|^{-s}$. Based on this observation, one might hope that Theorem 16.10 holds for all $s > 2$. This turns out to be true, and it was proven by Roth in the 1950s ([**Ro**]). Roth's proof builds on Thue's method but also introduces some important new ideas.

The theorems of Thue and Roth also say something about diophantine equations involving non-homogeneous polynomials. We explore this in the following exercise.

Exercise 16.4. Using Theorem 16.10, show that the following diophantine equation has only finitely many integer solutions:

$$y^9 - 2x^9 - 3x + y - 2 = 0.$$

Suppose $d \geq 3$. Suppose that $P_d(x, y)$ is a homogeneous polynomial of degree d with integer coefficients which is irreducible over $\mathbb{Z}$. Suppose that $Q(x, y)$ is a (not necessarily homogenenous) polynomial with integer coefficients and degree at most $d - 3$. Using Roth's theorem, show that the following diophantine equation has only finitely many integer solutions:

$$P_d(x, y) = Q(x, y).$$

16.4. Outline of Thue's proof

In this section we outline the proof of Theorem 16.10. We explain how the method extends the ideas from Liouville's proof. We also explain an analogy between Thue's proof and other arguments we have seen, such as the proof of finite field Nikodym.

Let us start by recalling the outline of Liouville's proof. Suppose that β is an algebraic number of degree d. By definition, this means that β is the root of a degree d polynomial Q with integer coefficients. If p/q is a rational number close to β, then by Taylor's theorem $|Q(p/q)| \lesssim |\beta - (p/q)|$. On the other hand, $Q(p/q)$ is a rational number with denominator q^d, and so it cannot be too small. Therefore, $|\beta - (p/q)|$ cannot be too small either. In rough terms, the polynomial Q "protects" β from rational approximations because $Q(\beta) = 0$ but $|Q(p/q)|$ cannot be too small.

Thue had the idea to use other polynomials besides Q to protect β. Other polynomials in one variable don't lead to any new estimate, but Thue had the remarkable idea to use polynomials in two variables. If $P(x_1, x_2) \in \mathbb{Z}[x_1, x_2]$ is a polynomial that vanishes (maybe along with some derivatives) at (β, β), then P can "protect" β from pairs of rational approximations $(p_1/q_1, p_2/q_2)$. To prove that there are only finitely many rational solutions to the inequality $|\beta - (p/q)| \leq |q|^{-2.6}$, this method requires infinitely many different auxiliary polynomials $P(x_1, x_2) \in \mathbb{Z}[x_1, x_2]$ which vanish at (β, β) to different orders. We will always make the convention that $q_2 > q_1$. It turns out that if q_2 has size roughly q_1^m, then it is helpful to use a polynomial P with degree 1 in the x_2 variable and degree roughly m in the x_1 variable, chosen so that $P(\beta, \beta) = 0$ and so that $\partial_1^j P(\beta, \beta) = 0$ for $1 \leq j \leq m - 1$.

Finding these auxiliary polynomials P is one of the interesting parts of the story. Thue carefully by hand crafted this infinite sequence of polynomials $P(x_1, x_2)$. He was able to construct the desired polynomials by hand when β is a d^{th} root of a rational number. He became stuck trying to generalize his method to other algebraic numbers, because he didn't know how to construct the auxiliary polynomials. At a certain point, Thue gave up trying to craft the polynomials he needed and instead, he proved that they must exist by counting parameters.

At the 1974 ICM, Schmidt gave a lecture [**Schm**] on Thue's work and its influence in number theory. He wrote,

The idea of asserting the existence of certain polynomials rather than explicitly con- structing them is the essential new idea in Thue's work. As Siegel [1970] points out, a study of Thue's papers reveals that Thue first tried hard to construct the polynomials explicitly (and he actually could do so in case β^d is rational).

This idea of finding an auxiliary polynomial P by counting parameters has a similar flavor to the parameter counting argument that we used in the proof of finite field Nikodym and throughout the book. Let us now review in outline the proof of finite field Nikodym and make a parallel outline of the proof of Theorem 16.10.

Outline of the proof of finite field Nikodym: Suppose that N is a small Nikodym set in $\mathbb{F}^n$.

(1) Find a non-zero polynomial P with controlled degree that vanishes on N. (Use parameter counting.)
(2) Because N is a Nikodym set, the polynomial P must also vanish at many other points. (Vanishing lemma.)
(3) The polynomial P vanishes at too many points, so it must be zero. Contradiction.

Here is the outline of the proof of Theorem 16.10. Suppose that the algebraic number β has two very good rational approximations $r_1 = p_1/q_1$ and $r_2 = p_2/q_2$.

(1) Find a non-zero polynomial $P \in \mathbb{Z}[x_1, x_2]$ with controlled degree and coefficients that vanishes to high order at (β, β). (Use parameter counting.)
(2) Because r_1 and r_2 are good approximations of β, the polynomial must also vanish to high order at (r_1, r_2).
(3) The polynomial P vanishes too much at (r_1, r_2), and so it must be zero. Contradiction.

We've talked a lot about the first step. To end this outline, let us say a little about steps 2 and 3.

Step 2 follows by Taylor's theorem. Since P vanishes to high order at (β, β), and since (r_1, r_2) is very close to (β, β), Taylor's theorem implies that $P(r_1, r_2)$ is very small. On the other hand, since P has integer coefficients, $P(r_1, r_2)$ is a rational number with denominator at most $q_1^{\mathrm{Deg}_{x_1} P} q_2^{\mathrm{Deg}_{x_2} P}$. So if $P(r_1, r_2)$ is sufficiently small, then $P(r_1, r_2)$ must be zero. The same argument applies to derivatives of P, and so we see that P and many derivatives vanish at (r_1, r_2).

Step 3 has to do with the size of the coefficients of P. In Step 1, we will pay attention to the size of the coefficients of P and get a good bound for them. If a polynomial with small coefficients vanishes at a rational number, then the denominator of the rational number cannot be too big. For polynomials in one variable, this is made precise by Gauss's lemma:

LEMMA 16.11. (Gauss) If $r = p/q$ and $P \in \mathbb{Z}[x]$ satisfies $\partial^j P(r) = 0$ for $j = 0, 1, \ldots, \ell - 1$, then $P(x) = (qx - p)^\ell P_1(x)$ for some $P_1 \in \mathbb{Z}[x]$. In particular, the leading coefficient of P has norm at least $|q|^\ell$.

We will recall the proof below and discuss a modification that applies to polynomials of two variables. Combining our bounds on the coefficients of P with this analysis, we will show that either q_1 or q_2 is bounded by $C(\beta)$.

This finishes the proof: it shows that β can have at most one very good rational approximation p/q with $q > C(\beta)$. It follows immediately that there are only finitely many very good rational approximations in total.

16.5. Step 1: Parameter counting

If L is a linear map from $\mathbb{R}^M$ to $\mathbb{R}^N$ with $M > N$, then there is a non-zero $x \in \mathbb{R}^M$ so that $Lx = 0$. We will need a variation of this result involving an integer solution x. If L is a linear map from $\mathbb{Z}^M$ to $\mathbb{Z}^N$, given by a matrix with integer entries, and if $M > N$, we will prove that there is a non-zero $x \in \mathbb{Z}^M$ so that $Lx = 0$. In fact, this follows from the pigeonhole principle by a very short argument.

We will also care about the size of the solution x. The argument above leads to a quantitative bound on the size of the entries of x. We let $|x|_\infty$ be the maximum size of any entry x_i of the vector x. Let $|L|_{op}$ be the operator norm of L with respect to the $|...|_\infty$ norms on $\mathbb{R}^M$ and $\mathbb{R}^N$. In other words, $|L|_{op}$ is the best constant in the inequality

$$|Lx|_\infty \leq |L|_{op}|x|_\infty.$$

LEMMA 16.12. (Siegel's lemma) If $L\colon \mathbb{Z}^M \to \mathbb{Z}^N$ is a linear map, given by a matrix with integer coefficients, with $M > N$, then there exists a nonzero $x \in \mathbb{Z}^M$ with $|x|_\infty \leq |L|_{op}^{N/(M-N)} + 1$ such that $Lx = 0$.

PROOF. For any $S \geq 0$, let us define $Q_S^M := \{x \in \mathbb{Z}^M : |x_i| \leq S, i = 1, \ldots, M\}$. By the definition of operator norm,

$$L : Q_S^M \to Q_{|L|_{op}S}^N.$$

Now $|Q_S^M| = (2S+1)^M$. We want to choose S so that the cardinality of the domain is greater than the cardinality of the range. In other words, we choose S so that

$$(2S+1)^M > (2|L|_{op}S + 1)^N. \tag{16.6}$$

For such a choice of S, the pigeonhole principle guarantees that there are $x_1, x_2 \in Q_S^M$ so that $Lx_1 = Lx_2$. But then $L(x_1 - x_2) = 0$. On the other hand, $|x_1 - x_2|_\infty \leq 2S$.

Now it just remains to pin down the size of S. This is just a little algebra. To get inequality 16.6, it suffices to have

$$(2S+1)^M > \left(|L|_{op}(2S+1)\right)^N,$$

which is equivalent to

$$2S + 1 > |L|_{op}^{N/(M-N)}.$$

We choose S to be the smallest integer obeying the last inequality, and so we get the bound

$$2S \leq |L|_{op}^{N/(M-N)} + 1.$$

□

Let us make a remark about the size of the solution x given by Siegel's lemma which will be useful for strategy later on. If $M = N+1$, then the Lemma guarantees an integer solution x, but the bound on $|x|_\infty$ is $\sim |L|_{op}^N$. As we increase M, we get much stronger bounds. For instance, if $M = (1.01)N$, then our bound on $|x|_\infty$ becomes $\sim |L|_{op}^{100}$.

Now we will use this parameter counting argument to find an integer polynomial $P(x_1, x_2)$ that vanishes at (β, β) to high order, with a bound on the degree of P and the size of the coefficients of P.

We write $|P|$ for the maximum of the norms of the coefficients of P.

PROPOSITION 16.13. Let $\beta \in \mathbb{R}$ be an algebraic number of degree d. Suppose $\varepsilon > 0$. For any integer sufficiently large integer m, there is a polynomial $P \in$

$\mathbb{Z}[x_1, x_2]$ with the form $P(x_1, x_2) = P_1(x_1)x_2 + P_0(x_1)$ and with the following properties:

- $\partial_1^j P(\beta, \beta) = 0$ for $0 \leq j \leq m-1$.
- $\text{Deg}\, P \leq (1+\epsilon)(1/2)dm + 2$.
- $|P| \leq C(\beta)^{m/\epsilon}$.

PROOF. The first step is to formulate this question so that finding P means finding an integer solution to a list of integer equations. Then we can apply Siegel's lemma.

We let D a degree to choose later. We write $P_1(x) = \sum_{i=0}^{D} b_i x^i$ and $P_0(x) = \sum_{i=0}^{D} a_i x^i$. The coefficients a_i and b_i define a point in $\mathbb{Z}^M$ with $M = 2D+2 > 2D$.

For each $0 \leq j \leq m-1$, we will write the equation $\partial_1^j P(\beta, \beta) = 0$ in terms of a_i and b_i. In order to do this, it's helpful to first note that

$$\partial_1^j x_1^i = i \cdot (i-1) \cdot ... \cdot (i-j+1) \cdot x^{i-j} = \frac{i!}{(i-j)!} x^{i-j}.$$

Plugging this in, for each $0 \leq j \leq m-1$, we get the equation

$$0 = \partial_1^j P(\beta, \beta) = \sum_i b_i \frac{i!}{(i-j)!} \beta^{i-j+1} + \sum_i a_i \frac{i!}{(i-j)!} \beta^{i-j}. \tag{16.7}$$

We now have a system of m equations for the a_i, b_i, with real coefficients. The coefficients are all in the field $\mathbb{Q}[\beta]$, but they are not integers. We will rewrite these m equations with $\mathbb{Q}[\beta]$ coefficients as a system of dm equations with integer coefficients.

Since β is an algebraic number of degree d, we know that it satisfies an equation of the form

$$Q(\beta) = \sum_{k=0}^{d} q_k \beta^k = 0,$$

where q_k are integers and $q_d \neq 0$. Therefore, β^d lies in the $\mathbb{Q}$-span of $1, \beta, ..., \beta^{d-1}$. On the other hand, $1, \beta, ..., \beta^{d-1}$ are linearly independent over $\mathbb{Q}$. Therefore, $1, \beta, ..., \beta^{d-1}$ form a $\mathbb{Q}$-basis for the field $\mathbb{Q}[\beta]$. In particular, for any exponent $e \geq d$, we can write

$$\beta^e = \sum_{k=0}^{d-1} C_{ke} \beta^k, \tag{16.8}$$

where C_{ke} are rational numbers. Plugging this expression into Equation 16.7, for each $0 \leq j \leq m-1$, we get an equation of the form

$$0 = \sum_{k=0}^{d-1} \beta^k \left[\sum_i b_i B_{ijk} + \sum_i a_i A_{ijk} \right] = 0,$$

where A_{ijk} and B_{ijk} are rational numbers. Since $1, \beta, ..., \beta^{d-1}$ are linearly independent over $\mathbb{Q}$, this equation is equivalent to the d equations

$$\sum_i b_i B_{ijk} + \sum_i a_i A_{ijk} = 0, \text{ for all } 0 \leq k \leq d-1. \tag{16.9}$$

After multiplying by a large constant to clear the denominators, we get d equations with integer coefficients for each value of j. In total, our original m

equations $\partial_1^j P(\beta,\beta) = 0$ for $j = 0, ..., m-1$ are equivalent to dm integer linear equations in the coefficients of P.

By Siegel's lemma, as long as $2D + 2 > dm$, we can find a non-zero integer polynomial $P(x_1, x_2) = P_1(x_1)x_2 + P_0(x_1)$ with $\mathrm{Deg}\, P_i \leq D$ so that $\partial_1^j P(\beta,\beta) = 0$ for $j = 0, 1, ..., m-1$. We can choose $D \leq (1/2)dm+1$, and so $\mathrm{Deg}\, P \leq (1/2)dm+2$. This polynomial P obeys the first two conditions in our Proposition.

It remains to estimate the size of the coefficients of P. To do this, we have to keep track of the size of the coefficients in Equation 16.7, and we have to estimate the size of the coefficients C_{ke} in Equation 16.8. It takes a little care and playing with the parameters to get a good bound. Recall that we want to prove the bound $|P| \leq C(\beta)^{m/\epsilon}$. It's worth mentioning at this point that a slightly weaker bound would not be good enough to prove Theorem 16.10 on diophantine approximation: for example a bound of the form m^m or C^{m^2} would be too big.

Let us return now to Equation 16.7. The coefficients $\frac{i!}{(i-j)!}$ can be quite big. The number i can be as large as $D \sim m$ and $j \leq m$, so this expression can be super-exponential in m. However, all the coefficients $\frac{i!}{(i-j)!}$ are divisible by $j!$, because

$$\frac{i!}{(i-j)!j!} = \binom{i}{j}.$$

Dividing through by $j!$, Equation 16.7 becomes:

$$0 = \frac{1}{j!}\partial_1^j P(\beta,\beta) = \sum_{i=0}^{D} b_i \binom{i}{j} \beta^{i-j+1} + \sum_{i=0}^{D} a_i \binom{i}{j} \beta^{i-j}. \tag{16.10}$$

We will choose $D \leq dm$, and so the coefficients $\binom{i}{j}$ obey the bound

$$\binom{i}{j} \leq 2^i \leq 2^D \leq C(\beta)^m.$$

Next we return to the expansion $\beta^e = \sum_{k=0}^{d-1} C_{ke}\beta^k$ and examine the size of the numerators and denominators of C_{ke}.

LEMMA 16.14. Suppose $Q(\beta) = 0$, where $Q \in \mathbb{Z}[x]$ with degree $\mathrm{Deg}(Q) = d$ and leading coefficient q_d. Then for any $e \geq d$, we can write

$$q_d^e \beta^e = \sum_{k=0}^{d-1} c_{ke}\beta^k,$$

where $c_{ke} \in \mathbb{Z}$ and $|c_{ke}| \leq [2|Q|]^e$.

PROOF. We have $0 = Q(\beta) = \sum_{k=0}^{d} q_k\beta^k$. We do the proof by induction on e, starting with $e = d$. For $e = d$, the equation $Q(\beta) = 0$ directly gives

$$q_d\beta^d = \sum_{k=0}^{d-1}(-q_k)\beta^k. \tag{$*$}$$

If we multiply both sides by q_d^{d-1}, we get a good expansion for the case $e = d$.

Now we proceed by induction. Suppose that $q_d^e\beta^e = \sum_{k=0}^{d-1} c_{ke}\beta^k$. Multiplying by $q_d\beta$, we get

$$q_d^{e+1}\beta^{e+1} = \sum_{k=0}^{d-1} c_{ke}q_d\beta^{k+1} = \sum_{k=1}^{d-1} c_{k-1,e}q_d\beta^k + \sum_{k=0}^{deg(\beta)-1} c_{d-1,e}(-q_k)\beta^k.$$

This formula shows how to write $c_{k,e+1}$ in terms of $c_{k,e}$. We see that $c_{k,e}$ are all integers and that $|c_{k,e+1}| \leq 2|Q| \max_k |c_{k,e}|$. This gives the required bound for $|c_{k,e}|$ by induction.

□

Now to convert Equation 16.10 into a set of d integer equations, we use Lemma 16.14 and then multiply by q_d^D. We get an equation of the form

$$0 = \sum_{k=0}^{d-1} \beta^k \left[\sum_i b_i B_{ijk} + \sum_i a_i A_{ijk} \right] = 0,$$

where A_{ijk} and B_{ijk} are integers of size at most $D2^D[2|Q|]^D \leq C(\beta)^m$.

To summarize, we have now converted the equations $\partial_1^j P(\beta, \beta) = 0$ for $j = 0, ..., m-1$ into a system of md integer equations with coefficients of size at most $C(\beta)^m$.

We now apply Siegel's lemma. The operator norm of the linear operator L is at most $C(\beta)^m$. The domain has dimension $M = 2D + 2$. The target has dimension $N = dm$. If we choose D so that $M = N + 1$ or $N + 2$, then $|P|$ will turn out to be too large. This is the purpose of the ε in the degree bound for P. Since m is assumed to be large enough, we can can choose D so that $D \leq (1+\varepsilon)(1/2)dm$, and yet

$$M = 2D + 2 \geq (1 + \frac{\varepsilon}{2})dm = (1 + \frac{\varepsilon}{2})N.$$

Using Siegel's lemma, it now follows that we can find a solution P with

$$|P| \leq 1 + (C(\beta)^m)^{\frac{N}{M-N}} \leq 1 + C(\beta)^{2m/\varepsilon}.$$

This gives the desired bound on $|P|$ and finishes the proof of the proposition. □

16.6. Step 2: Taylor approximation

PROPOSITION 16.15. Suppose that β is an algebraic number of degree $d \geq 3$. Suppose that $s > \frac{d+2}{2}$. There is a small constant $c(\beta, s) > 0$ so that the following holds.

Suppose that $r_1 = p_1/q_1$ and $r_2 = p_2/q_2$ obey

$$|\beta - r_i| \leq q_i^{-s}.$$

We assume that $q_1 < q_2$, and we let m be the integer so that

$$q_1^m \leq q_2 < q_1^{m+1}.$$

Given β and s, we also assume that q_1 is sufficiently large and that m is sufficiently large.

Then there exists a polynomial $P \in \mathbb{Z}[x_1, x_2]$, $P(x_1, x_2) = P_1(x_1)x_2 + P_0(x_1)$, so that

- $\partial_1^j P(r_1, r_2) = 0$ for $0 \leq j < c(\beta, s)m$.
- $|P| \leq C(\beta, s)^m$.
- $\operatorname{Deg} P \leq C(\beta)m$.

PROOF. We let P be the polynomial given by Proposition 16.13. In this Proposition, we need to choose an ε, and we take

$$\varepsilon = (1/10d)(s - \frac{d+2}{2}).$$

Proposition 16.13 immediately gives us items 2 and 3 above, and it also tells us that $\partial_1^j P(\beta,\beta) = 0$ for $0 \le j \le m-1$. It just remains to check that $\partial_1^j P(r_1,r_2) = 0$ for $0 \le j \le c(\beta,s)m$.

To prove this, we will use Taylor's theorem to estimate $|P(r_1,r_2)|$. We recall Taylor's theorem.

THEOREM 16.16. If f is a smooth function on an interval, then $f(x+h)$ can be approximated by its Taylor expansion around x:
$f(x+h) = \sum_{j=0}^{m-1}(1/j!)\partial^j f(x)h^j + E$,
where the error term E is bounded by
$|E| \le (1/m!)\sup_{y\in[x,x+h]}|\partial^m f(y)|h^m$.

In particular, if f vanishes to high order at x, then $f(x+h)$ will be very close to $f(x)$.

COROLLARY 16.17. If Q is a polynomial of one variable, and Q vanishes at x to order $m \ge 1$, and if $|h| \le 1$, then

$$|Q(x+h)| \le C(x)^{\operatorname{Deg} Q}|Q|h^m.$$

PROOF. We have to estimate the size of the coefficients of $(1/m!)\partial^m Q$. We recall that

$$(1/m!)\partial_x^m x^i = \binom{i}{m}x^{i-m},$$

and so the coefficients of $(1/m!)\partial^m Q$ have norm $\le 2^{\operatorname{Deg} Q}|Q|$.

Therefore, we get

$$\sup_{|y-x|\le 1}(1/m!)|\partial^m Q(y)| \le 2^{\operatorname{Deg} Q}|Q|(\operatorname{Deg} Q)(|x|+1)^{\operatorname{Deg} Q} \le C(x)^{\operatorname{Deg} Q}|Q|.$$

Plugging this estimate into Taylor's theorem finishes the proof. □

We want to use this estimate to bound $|P(r_1,r_2)|$. More generally we also want to bound $|\partial_1^j P(r_1,r_2)|$ for the given range of $j : 0 \le j \le c(\beta,s)m$. We will do all these estimates at the same time. We fix a value of j in the range $0 \le j \le c(\beta,s)m$, and we define

$$\tilde{P}(x_1,x_2) = (1/j!)\partial_1^j P(x_1,x_2).$$

We note that $\tilde{P}$ still has integer coefficients. Clearly $\operatorname{Deg}\tilde{P} \le \operatorname{Deg} P$. Also $|\tilde{P}|$ obeys essentially the same bound as $|P|$: $|\tilde{P}| \le 2^{\operatorname{Deg} P}|P| \le C(\beta,s)^m$.

Let $Q(x) = \tilde{P}(x,\beta)$. The polynomial Q vanishes to order $(1-c(\beta,s))m$ at $x=\beta$, and $|Q| \le C(\beta,s)^m$.

From the corollary we see that

$$|\tilde{P}(r_1,\beta)| \le C(\beta,s)^m|\beta - r_1|^{(1-c(\beta,s))m}.$$

On the other hand, $\partial_2\tilde{P}$ is bounded by $C(\beta,s)^m$ in a unit disk around (β,β), and so

$$|\tilde{P}(r_1,r_2) - \tilde{P}(r_1,\beta)| \le C(\beta,s)^m|\beta - r_2|.$$

Combining these, we see that

$$|\tilde{P}(r_1,r_2)| \le C(\beta,s)^m\left[|\beta-r_1|^{(1-c(\beta,s))m} + |\beta - r_2|\right].$$

Now we can use the assumption that $|\beta - r_i| \le q_i^{-s}$.

$$|\tilde{P}(r_1,r_2)| \le C(\beta,s)^m\left[q_1^{-s(1-c)m} + q_2^{-s}\right].$$

We note that $q_1^m \leq q_2$, and so $q_2^{-s} \leq q_1^{-ms}$, and so the second term is dominated by the first term. Therefore, we get

$$|\tilde{P}(r_1, r_2)| \leq C(\beta, s)^m q_1^{-(1-c)sm}. \tag{16.11}$$

On the other hand, $\tilde{P}$ has integer coefficients, and so $\tilde{P}(r_1, r_2)$ is a rational number with denominator at most $q_1^{\mathrm{Deg}_{x_1} P} q_2$. By Proposition 16.13, we know that $\mathrm{Deg}\, P \leq (1+\varepsilon)(1/2)dm$. Also, we know that $q_2 \leq q_1^{m+1}$. Therefore, we the denominator of $P(r_1, r_2)$ is at most

$$q_1^{(1+\varepsilon)(d/2)m+m+1}. \tag{16.12}$$

We now claim that $|\tilde{P}(r_1, r_2)|$ is so small that it must vanish. Since we have assumed that q_1 is very large compared to $C(\beta, s)$, we just have to check that the exponents of q_1 in Equation 16.11 is more extreme than the exponent in Equation 16.12. In other words, we have to check that

$$(1-c)sm > (1+\varepsilon)(d/2)m + m + 1.$$

By our definition of ε, $s = \frac{d+2}{2} + 10d\varepsilon$. And so the last equation is equivalent to

$$(1-c)(\frac{d}{2} + 1 + 10d\varepsilon)m > (\frac{d}{2} + 1 + \frac{d}{2}\varepsilon)m + 1.$$

Since we assumed m is large, the final $+1$ is negligible, and the resulting equation is true as long as we pick $c = c(\beta, s)$ sufficiently small comported to $\varepsilon = \varepsilon(\beta, s)$. □

16.7. Step 3: Gauss's lemma

To finish the proof we want to show that an integer polynomial with small coefficients cannot vanish at a rational point with a large denominator. We begin by considering by polynomials of one variable, where Gauss's lemma gives a sharp estimate.

LEMMA 16.18. (Gauss's lemma) If $r = p/q$ is a rational number, and if $P \in \mathbb{Z}[x]$ satisfies $\partial^j P(r) = 0$ for $j = 0, 1, \ldots, l-1$, then $P(x) = (qx - p)^l P_1(x)$ for some $P_1 \in \mathbb{Z}[x]$.

PROOF. The vanishing condition tells us that $P(x) = (qx-p)^l P_2(x)$ for some polynomial $P_2 \in \mathbb{R}[x]$. It remains to show that the coefficients of P_2 are integers. By expanding the equation $P(x) = (qx-p)^l P_2(x)$, and solving for the coefficients of P_2, we see that the coefficients of P_2 must be rational.

Taking out the lowest common denominator, we write $P(x) = \frac{1}{M}(qx-p)^l \tilde{P}_2(x)$, for some $\tilde{P}_2 \in \mathbb{Z}[x]$ so that there is no prime dividing all the coefficients of $\tilde{P}_2$ as well as M. So $MP(x) = (qx-p)^l \tilde{P}_2(x)$. If $M \neq \pm 1$, then let s be any prime divisor of M. Then we get a contradiction modulo s, since $qx-p$ is not 0 mod s as p/q was already given in lowest terms, and $\tilde{P}_2$ is also not 0 mod s. It follows that $M = \pm 1$ and hence $P_1 \in \mathbb{Z}[x]$. □

As a corollary, we see that if $P \in \mathbb{Z}[x]$ vanishes to order l at a rational point p/q, then q^l divides the leading coefficient of P, and so $|P| \geq q^l$.

Now we adapt these ideas to polynomials in two variables.

PROPOSITION 16.19. If $P(x_1, x_2) = P_1(x_1)x_2 + P_0(x_1) \in \mathbb{Z}[x_1, x_2]$, and $(r_1, r_2) = (p_1/q_1, p_2/q_2) \in \mathbb{Q}^2$, and $\partial_1^j P(r_1, r_2) = 0$ for $j = 0, ..., l-1$, and if $l \geq 2$, then

$$|P| \geq \min((2DegP)^{-1} q_1^{\frac{l-1}{2}}, q_2).$$

Remark. We need to assume that $l \geq 2$ to get any estimate. For instance, the polynomial $P(x_1, x_2) = 2x_1 - x_2$, vanishes at $(r_1, 2r_1)$ for any rational number r_1. But $|P| = 2$, and q_1 can be arbitrarily large.

As soon as $l \geq 2$, the size of $|P|$ constrains the complexity of the rational point (r_1, r_2). It can still happen that one component of r is very complicated, but they can't both be very complicated.

PROOF. Our assumption is that

$$\partial^j P_1(r_1) r_2 + \partial^j P_0(r_1) = 0, 0 \leq j \leq l-1.$$

Let $V(x)$ be the vector $(P_1(x), P_0(x))$. Our assumption is that for $0 \leq j \leq l-1$, the derivatives $\partial^j V(r_1)$ all lie on the line $V \cdot (r_2, 1) = 0$. In particular, any two of these derivatives are linearly dependent. This tells us that many determinants vanish. If V and W are two vectors in $\mathbb{R}^2$, we write $[V, W]$ for the 2×2 matrix with first column V and second column W. Therefore,

$$\det[\partial^{j_1} V, \partial^{j_2} V](r_1) = 0, \text{ for any } 0 \leq j_1, j_2 \leq l-1.$$

Because the determinant is multilinear, we have the Leibniz rule $\partial \det[V, W] = \det[\partial V, W] + \det[V, \partial W]$, which holds for any vector-valued functions $V, W : \mathbb{R} \to \mathbb{R}^2$. Using this Leibniz rule, we see that

$$\partial_j \det[V, \partial V](r_1) = 0, \text{ for any } 0 \leq j \leq l-2.$$

Now $\det[V, \partial V]$ is a polynomial in one variable with integer coefficients. If this polynomial is non-zero, then by Gauss's lemma we conclude that

$$|\det[V, \partial V]| \geq q_1^{l-1}.$$

Expanding out in terms of P, we have $|\det[V, \partial V]| = |\partial P_0 P_1 - \partial P_1 P_0| \leq 2(DegP)^2 |P|^2$. Therefore, we have $|P| \geq (2DegP)^{-1} q_1^{\frac{l-1}{2}}$.

The polynomial $\det[V, \partial V]$ may also be identically zero. In this case, the polynomial P must simplify dramatically. One possibility is that P_1 is identically zero. In this case $P(x_1, x_2) = P_0(x_1)$, and by the Gauss lemma we have that $|P| \geq q_1^l$. If P_1 is not identically zero, then the derivative of the ratio P_0/P_1 is identically zero. (The numerator of this derivative is $\det[V, \partial V]$.) This implies that $P_0 = AP_1$ for some number A, and so

$$P(x_1, x_2) = (x_2 + A) P_1(x_1).$$

This scenario leaves two possibilities. Either $r_2 + A = 0$ or else, $\partial_1^j P_1(r_1) = 0$ for $0 \leq j \leq l-1$. In the second case, Gauss's lemma implies that $|P_1| \geq q_1^l$. In the first case $A = -r_2 = -p_2/q_2$. In this case, we can write P in the form

$$P(x_1, x_2) = (q_2 x_2 - p_2) \tilde{P}(x_1),$$

where a priori $\tilde{P}(x_1)$ has rational coefficients. However, the same argument as in the proof of Gauss's lemma shows that $\tilde{P}$ actually has integer coefficients. Therefore, $|P| \geq q_2$. □

16.8. Conclusion

We now have all the tools to quickly prove Theorem 16.10. Suppose that β is an algebraic number of degree d and that $s > \frac{d+2}{2}$. We have to show that there are only finitely many rational solutions to the inequality

$$|\beta - \frac{p}{q}| \leq q^{-s}.$$

We give a proof by contradiction. Suppose that there are infinitely many such rational numbers p/q. Let p_1/q_1 be one rational solution, where q_1 is extremely large. Then let p_2/q_2 be another rational solution, with q_2 much larger than q_1. We define m to be the integer so that $q_1^m \leq q_2 < q_1^{m+1}$.

By Proposition 16.15, there is a polynomial $P(x_1, x_2) \in \mathbb{Z}[x_1, x_2]$ of the form $P(x_1, x_2) = P_1(x_1)x_2 + P_0(x_1)$ so that

- $\partial_1^j P(r_1, r_2) = 0$ for $0 \leq j \leq l-1$, with $l = c(\beta, s)m$.
- $|P| \leq C(\beta, s)^m$.
- $\operatorname{Deg} P \lesssim m$.

On the other hand, Proposition 16.19 gives a lower bound for $|P|$:

$$|P| \geq \min\left(m^{-1} q_1^{\frac{l-1}{2}}, q_2\right) \geq m^{-1} q_1^{c(\beta,s)m}.$$

Comparing the bounds, we see that

$$m^{-1} q_1^{c(\beta,s)m} \leq C(\beta, s)^m,$$

and so

$$q_1 \leq C(\beta, s).$$

Since we were allowed to choose q_1 arbitrarily large, this gives a contradiction.

Exercise 16.5. In this exercise, we estimate how sharp Proposition 16.19 is. Suppose that $P \in \mathbb{Z}[x_1, x_2]$ has the form

$$P(x_1, x_2) = P_1(x_1)x_2 + P_0(x_1).$$

Suppose that $r = (r_1, r_2) \in \mathbb{Q}^2$, and suppose that $\partial_1^j P(r) = 0$ for $0 \leq j \leq l-1$. How small can $|P|$ be?

Here are two explicit examples: the polynomial $q_2 x_2 - p_2$ which has $|P| \geq q_2$, and the polynomial $(q_1 x_1 - p_1)^l$, which has $|P| \geq q_1^l$.

If l is large and $q_2 \sim q_1^{l/2}$, and if $\operatorname{Deg} P$ is on the order of l, then the lower bound from Proposition 16.19 is approximately $|P| \gtrsim q_1^{l/2}$. The lower bound is much smaller than the two explicit examples.

Using parameter counting, prove the following Proposition which shows that the bound from Proposition 16.19 is fairly sharp.

Proposition 16.20. For any $r \in \mathbb{Q}^2$, and any $l \geq 0$, $\varepsilon > 0$, there is a polynomial $P \in \mathbb{Z}[x_1, x_2]$ with the form $P(x_1, x_2) = P_1(x_1)x_2 + P_0(x_1)$ obeying the following conditions.

- $\partial_1^j P(r) = 0$ for $j = 0, ..., l-1$.
- $|P| \leq C(\epsilon)^l (p_1 + q_1)^{\frac{l}{2}+\epsilon}$.
- The degree of P is $\lesssim \epsilon^{-1}\left(l + \log_{\|r_1\|} \|r_2\|\right)$.

Exercise 16.6. Suppose that α and β are two algebraic numbers. Prove that $\alpha + \beta$ is an algebraic number.

EXERCISE 16.7. Using Gauss's lemma, check the following. Suppose that Q is a polynomial with integer coefficients. If Q vanishes at a rational point p/q, then $Q(x) = (qx - p)Q_1(x)$ where Q_1 has integer coefficients. Therefore, if β is an algebraic number and Q is an integer polynomial of minimal degree vanishing at β, then Q does not have any rational roots.

EXERCISE 16.8. Suppose that $Q \in \mathbb{Z}[x]$ is irreducible over $\mathbb{Z}$, and that $\beta \in \mathbb{C}$ is a root of Q. Then $Q'(\beta) \neq 0$. In other words, every root of Q has multiplicity 1.

Suppose that $\beta = \beta_1, \beta_2, ..., \beta_d$ are the Galois conjugates of β in $\bar{\mathbb{Q}}$. Consider the polynomial

$$R(x) = \prod_{j=1}^{d} (x - \beta_j).$$

Any symmetric function of the β_j is invariant under the action of the Galois group, and so any symmetric function of the β_j is rational. Therefore, $R(x) \in \mathbb{Q}[x]$. We multiply R by an integer C to get a polynomial $R_1(x) \in \mathbb{Z}[x]$, and we choose C so that the gcd of the coefficients of R_1 is 1. Note that R_1 vanishes at each β_j with multiplicity 1.

Since β_j are Galois conjugates of β, we see that $Q(\beta_j) = 0$ for all j, and so $R_1(x)$ divides $Q(x)$ in the ring $\mathbb{Q}[x]$.

Imitating the proof of Lemma 16.18, show that $Q(x) = R_1(x)Q_1(x)$ where $Q_1 \in \mathbb{Z}[x]$. Since Q is irreducible, $Q_1(x)$ must be a constant. Therefore $Q(x)$ vanishes at each β_j with multiplicity 1. In particular, $Q'(\beta) \neq 0$.

EXERCISE 16.9. In this exercise, we describe another class of examples of diophantine equations $P(x, y) = 0$ with infinitely many integer solutions, generalizing the examples at the end of Section 16.1.

Let $\Gamma \subset \mathbb{R}^2$ be the curve $y^2 - 2x^2 = 1$. By Proposition 16.2, Γ contains infinitely many integer points. Now consider a map $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ where the components, ϕ_1 and ϕ_2, are polynomials with integer coefficients. The map ϕ sends integer points to integer points. For many choices of ϕ, the map $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ is finite-to-one, and so the image of ϕ contains infinitely many integer points.

Prove that the image of ϕ lies in a curve $Z(P)$, where $P(x, y)$ is a polynomial with integer coefficients. Give examples where P is irreducible and has arbitrarily high degree.

There is a beautiful theorem of Siegel that classifies when a diophantine equation of two variables, $P(x, y) = 0$, has infinitely many integer solutions. It is a little beyond the scope of this book to state Siegel's theorem, but the rough idea is that if $P(x, y) = 0$ has infinitely many solutions, then they must come from a parametrization like the one at the end of Section 16.1 or the one in Exercise 16.9.

Bibliography

[AA] P. Agarwal, B. Aronov, Counting facets and incidences, Discrete Comput. Geom. 7 (1992) 359-369.

[ACNS] M. Ajtai, V. Chvátal, M. Newborn, and E. Szemerédi, Crossing-free subgraphs. Theory and practice of combinatorics, 9-12, North-Holland Math. Stud., 60, North-Holland, Amsterdam, 1982.

[AjSz] M. Ajtai and E. Szemerédi, Sets of Lattice Points That Form No Squares, *Studia. Scientiarum Mathematicarum Hungarica.* **9** (1974), 9-11.

[Al] N. Alon, Combinatorial nullstellensatz, Recent trends in combinatorics (Matrahaza, 1995). Combin. Probab. Comput. 8 (1999), no. 1-2, 7-29.

[AlSp] N. Alon and J. Spencer, *The Probabilistic Method* Third edition. With an appendix on the life and work of Paul Erd?s. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley and Sons, Inc., Hoboken, NJ, 2008.

[ApSh] R. Apfelbaum and M. Sharir, On incidences between points and hyperplanes,

[Ar] V. I. Arnold, The principle of topological economy in algebraic geometry. Surveys in modern mathematics, 13-23, London Math. Soc. Lecture Note Ser., 321, Cambridge Univ. Press, Cambridge, 2005.

[ACNS] M. Ajtai, V. Chvatal, M. Newborn, and E. Szemerédi, Crossing-free subgraphs. Theory and practice of combinatorics, 9-12, North-Holland Math. Stud., 60, North-Holland, Amsterdam, 1982.

[ArSh] B. Aronov, and M. Sharir, Cutting circles into pseudo-segments and improved bounds for incidences. Discrete Comput. Geom. 28 (2002), no. 4, 475-490.

[ALMSS] Arora, Sanjeev; Lund, Carsten; Motwani, Rajeev; Sudan, Madhu; Szegedy, Mario, Proof verification and the hardness of approximation problems. J. ACM 45 (1998), no. 3, 501-555.

[AS] S. Arora and S. Safra, Probabilistic checking of proofs: a new characterization of NP. J. ACM 45 (1998), no. 1, 70-122.

[BF] L. Babai and P. Frankl, *Linear Algebra Methods in Combinatorics.*

[BCT] J. Bennett, A. Carbery, T. Tao, On the multilinear restriction and Kakeya conjectures. Acta Math. 196 (2006), no. 2, 261-302.

[BW] E. Berlekamp and L. Welch, Error correction of algebraic block codes. US Patent Number 4,633,470. 1986.

[BC] R. C. Bose and I. M. Chakravarti, Hermitian varieties in a finite projective space $PG(N, q^2)$, Canad. J. Math. 18 (1966), 1161-1182.

[BKT] J. Bourgain, N. Katz, T. Tao, A sum-product estimate in finite fields, and applications. Geom. Funct. Anal. 14 (2004), no. 1, 27-57.

[Bou] J. Bourgain, On the Erdős-Volkmann and Katz-Tao ring conjectures. Geom. Funct. Anal. 13 (2003), no. 2, 334-365.

[BrKn] P. Brass and C. Knauer, On counting point-hyperplane incidences, Comput. Geom. Theory Appl. 25 (1-2) (2003) 13-20.

[BrSc] A. Brouwer, and A. Schrijver, The blocking number of an affine space. J. Combinatorial Theory Ser. A 24 (1978), no. 2, 251-253.

[CS] L. Carleson and P. Sjolin, Oscillatory integrals and a multiplier problem for the disc. Collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity, III. Studia Math. 44 (1972), 287-299.

[CEGPSSS] B. Chazelle, H. Edelsbrunner, L. Guibas,R. Pollack, R.Seidel, M. Sharir, and J. Snoeyink, Counting and cutting cycles of lines and rods in space, Computational Geometry: Theory and Applications, 1(6) 305-323 (1992).

[CEGSW] K.L. Clarkson, H. Edelsbrunner, L. Guibas, M Sharir, and E. Welzl, Combinatorial Complexity bounds for arrangements of curves and spheres, Discrete Comput. Geom. (1990) 5, 99-160.

[CrLe] E. Croot, and V. F. Lev, Problems presented at the Workshop on Recent Trends in Additive Combinatorics, 2004, American Institute of Mathematics, Palo Alto, CA.

[Ca] M. do Carmo, *Differential geometry of curves and surfaces*, Translated from the Portuguese. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1976.

[Ch] B. Chazelle, Cutting hyperplanes for divide-and-conquer, Discrete Comput. Geom. 9 (1993) 145-158.

[CL] S.S. Chern and R.K. Lashof, On the Total Curvature of Immersed Manifolds, American Journal of Mathematics, 79, (1957) No. 2., 306-318.

[D] Z. Dvir, On the size of Kakeya sets in finite fields, J. Amer. Math Soc. (2009) 22, 1093-1097.

[D2] Z. Dvir, Incidence theorems and their applications. Found. Trends Theor. Comput. Sci. 6 (2010), no. 4, 257-393 (2012).

[EGS] H. Edelsbrunner, L. Guibas, and M. Sharir, The complexity of many cells in arrangements of planes and related problems, Discrete Comput. Geom. (1990) 5, 197-216.

[Ei] D. Eisenbud, *Commutative Algebra with a View Toward Algebraic Geometry*, Springer, Graduate Texts in Mathematics 150, 2004.

[Er1] P. Erdős, On sets of distances of n points, Amer. Math. Monthly (1946) 53, 248-250.

[Er2] P. Erdős, Some of my favorite problems and results, in *The Mathematics of Paul Erdős*, Springer, 1996.

[El1] Gy. Elekes, n points in the plane can determine $n^{3/2}$ unit circles. Combinatorica 4 (1984), no. 2-3, 131.

[El2] Gy. Elekes, SUMS versus PRODUCTS in number theory, algebra and Erd?s geometry. Paul Erdős and his mathematics, II (Budapest, 1999), 241-290, Bolyai Soc. Math. Stud., 11, Janos Bolyai Math. Soc., Budapest, 2002.

[EKS] Gy. Elekes, H. Kaplan, and M. Sharir, On lines, joints, and incidences in three dimensions, Journal of Combinatorial Theory, Series A (2011) 118, 962-977.

[ElSh] Gy. Elekes and M. Sharir, Incidences in three dimensions and distinct distances in the plane, Proceedings 26th ACM Symposium on Computational Geometry (2010) 413-422.

[ESS] Gy. Elekes, M. Simonovits, and E. Szabó, A combinatorial distinction between unit circles and straight lines: how many coincidences can they have? Combin. Probab. Comput. 18 (2009), no. 5, 691-705.

[ElTo] Gy. Elekes and C. Toth, Incidences of not-too-degenerate hyperplanes. Computational geometry (SCG'05), 16-21, ACM, New York, 2005.

[ElHa] J. Ellenberg and M. Hablicsek, An incidence conjecture of Bourgain over fields of positive characteristic, arXiv:1311.1479

[Fed] H. Federer, *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153 Springer-Verlag New York Inc., New York 1969.

[Fef] C. Fefferman, The multiplier problem for the ball, Ann. of Math. (2) 94 (1971), 330-336.

[Fef2] C. Fefferman, Inequalities for strongly singular convolution operators. Acta Math. 124 1970 9-36.

[FS] S. Feldman and M. Sharir, *An improved bound for joints in arrangements of lines in space*, Discrete Comput. Geom. (2005) 33, 307-320.

[Ful] W. Fulton, *Algebraic Topology, a First Course*, Springer-Verlag, Graduate Texts in Mathematics 153, 1995.

[GT] B. Green and T. Tao, On sets defining few ordinary lines. Discrete Comput. Geom. 50 (2013), no. 2, 409-468.

[Gr] , M. Gromov, Dimension, non-linear spectra and width, Lect. Notes in Math. Springer-Verlag 1317 (1988), 132-185.

[Gr2] M. Gromov, Isoperimetry of waists and concentration of maps. Geom. Funct. Anal. 13 (2003), no. 1, 178-215.

[GP] V. Guillemin and A. Pollack, *Differential Topology* Reprint of the 1974 original. AMS Chelsea Publishing, Providence, RI, 2010.

[Gu1] L. Guth, The width-volume inequality. Geom. Funct. Anal. 17 (2007), no. 4, 1139-1179.

[Gu2] Minimax problems related to cup powers and Steenrod squares. Geom. Funct. Anal. 18 (2009), no. 6, 1917-1987.

[Gu3] L. Guth, The endpoint case of the Bennett-Carbery-Tao multilinear Kakeya conjecture. Acta Math. 205 (2010), no. 2, 263-286.

[Gu4] L. Guth, Distinct distance estimates and low degree polynomial partitioning. Discrete Comput. Geom. 53 (2015), no. 2, 428-444.

[Gu5] L. Guth, A restriction estimate based on polynomial partitioning. arXiv:1407.1916

[Gu6] L. Guth, Polynomial methods in combinatorics and Fourier analysis, 2015 Namboodiri lectures at the University of Chicago, Notes available on Guth's webpage: http://math.mit.edu/ lguth/

[GK1] L. Guth and N. Katz, Algebraic methods in discrete analogs of the Kakeya problem. Adv. Math. 225 (2010), no. 5, 2828-2839.

[GK2] L. Guth and N. Katz, On the Erdős distinct distance problem in the plane, Annals of Math 181 (2015), 155-190.

[GS] L. Guth and A. Suk, The joints problem for matroids. J. Combin. Theory Ser. A 131 (2015), 71-87.

[HL] R. Harvey, and H. B. Lawson, Calibrated geometries. Acta Math. 148 (1982), 47-157.

[Ha] J. Harris, *Algebraic Geometry, a first course*, Corrected reprint of the 1992 original. Graduate Texts in Mathematics, 133. Springer-Verlag, New York, 1995.

[Ja] V. Jarnik, Uber die Gitterpunkte auf konvexen Kurven. (German) Math. Z. 24 (1926), no. 1, 500-518.

[JP] R. Jerrard and M. Pakzad, Sobolev spaces of isometric immersions of arbitrary dimension and codimension, arXiv:1405.4765.

[Ka] E. Kaltofen, Polynomial factorization 1987-1991 in "Latin '92," (I. Simon, Ed.) Lecture Notes in Computer Science, Vol. 585, 294-313, Springer-Verlag, New York/Berlin, 1992.

[KMS] H. Kaplan, J. Matoušek, and M. Sharir, Simple proofs of classical theorems in discrete geometry via the Guth–Katz polynomial partitioning technique. Discrete Comput. Geom. 48 (2012), no. 3, 499-517.

[KSS] H. Kaplan, M. Sharir, and E. Shustin, On lines and joints, Discrete Comput Geom (2010) 44, 838-843.

[Ka] N. Katz, The flecnode polynomial: a central object in incidence geometry, Proceedings of the 2014 ICM, arXiv:1404.3412

[KLT] N. Katz, I. Laba, and T. Tao, An improved bound on the Minkowski dimension of Besicovitch sets in R^3, Ann. of Math. (2) 152 (2000), no. 2, 383-446.

[KT1] N. Katz and T. Tao, New bounds for Kakeya problems. Dedicated to the memory of Thomas H. Wolff. J. Anal. Math. 87 (2002), 231-263.

[KT2] N. Katz and T. Tao, Some connections between Falconer's distance set conjecture and sets of Furstenburg type. New York J. Math. 7 (2001), 149-187.

[KatTar] N. Katz and G. Tardos, A new entropy inequality for the Erdős distance problem, Towards a theory of geometric graphs, 119-126, Contemp. Math., 342, Amer. Math. Soc., Providence, RI, 2004.

[Ko] J. Kollár, Szemer di-Trotter-type theorems in dimension 3. Adv. Math. 271 (2015), 30-61.

[KRS] J. Kollár, L. Rónyai, and T. Szabó, Norm-graphs and bipartite Turin numbers. Combinatorica 16 (1996), no. 3, 399-406.

[Koo] T. Koornwinder, A note on the absolute bound for systems of lines, Proc. Konink. Nedrl. Akad. Wet. Ser A, 79 (1976) 152-3.

[KM] P. Kronheimer and T. Mrowka, Gauge theory for embedded surfaces. I. Topology 32 (1993), no. 4, 773-826.

[Lak] I. Lakatos, *Proofs and refutations* The logic of mathematical discovery. Edited by John Worrall and Elie Zahar. Cambridge University Press, Cambridge-New York-Melbourne, 1976.

[Lab] I. Laba, From harmonic analysis to arithmetic combinatorics. Bull. Amer. Math. Soc. (N.S.) 45 (2008), no. 1, 77-115.

[Lan] S. Lang, *Algebra*, Revised Third Edition, Springer, Graduate Texts in Mathematics 211, New York 2002.

[LRS77] D. Larman, C. Rogers, J. Seidel, On two-distance sets in Euclidean space. Bull. London Math. Soc. 9 (1977) 261-7.

[Le] F. T. Leighton, (1983) Complexity Issues in VLSI. MIT Press.

[LW] L. Loomis and H. Whitney, An inequality related to the isoperimetric inequality. Bull. Amer. Math. Soc 55, (1949). 961-962.

[LuSu] A. Guo, S. Kopparty, and M. Sudan, New affine-invariant codes from lifting, arXiv:1208.5413

[Ma] J. Matoušek, *Using the Borsuk-Ulam theorem, Lectures on topological methods in combinatorics and geometry.* Written in cooperation with Anders Bjőrner and Gunter M. Ziegler. Universitext. Springer-Verlag, Berlin, 2003.

[Ma2] J. Matoušek, *Thirty three miniatures* Mathematical and algorithmic applications of linear algebra. Student Mathematical Library, 53. American Mathematical Society, Providence, RI, 2010.

[Mi] J. Milnor, On the Betti numbers of real varieties. Proc. Amer. Math. Soc. 15 1964 275-280.

[MT] G. Mockenhaupt and T. Tao, Restriction and Kakeya phenomena for finite fields. Duke Math. J. 121 (2004), no. 1, 35-74.

[NW] Z. Nie and A. Wang, Hilbert functions and the finite degree Zariski closure in finite field combinatorial geometry. J. Combin. Theory Ser. A 134 (2015), 196-220.

[PS] J. Pach and M. Sharir, Combinatorial geometry and its algorithmic applications. The Alcalá lectures. Mathematical Surveys and Monographs, 152. American Mathematical Society, Providence, RI, 2009.

[Q] R. Quilodrán, The joints problem in $\mathbf{R^n}$, Siam J. Discrete Math, Vol. 23, 4, p. 2211-2213.

[Re] Reiman, "Uber ein Problem von K. Zarankiewicz", Acta Mathematica Hungarica 9: 269-273.

[Ro] K. Roth, Rational approximations to algebraic numbers. Mathematika 2 (1955), 1-20.

[Sa] L. Santalo, Integral geometry and geometric probability. Second edition. With a foreword by Mark Kac. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2004.

[Schm] W. Schmidt, Applications of Thue's method in various branches of number theory. Proceedings of the International Congress of Mathematicians (Vancouver, B.C., 1974), Vol. 1, pp. 177-185. Canad. Math. Congress, Montreal, Que., 1975.

[Sch] I. J. Schoenberg. On the Besicovitch-Perron solution of the Kakeya problem. In Studies in mathematical analysis and related topics, pages 359-363. Stanford Univ. Press, Stanford, Calif., 1962

[Seg] B. Segre, Forme e geometrie hermitiane, con particolare riguardo al caso finito, Ann. Mat. Pura Appl. (4) 70 (1965), 1-201.

[SSS] M. Sharir, A. Sheffer, and J. Solymosi, Distinct distances on two lines, J. Combin. Theory Ser. A 120 (2013), no. 7, 1732-1736.

[SoSt] J. Solymosi and M. Stojaković, Many collinear k-tuples with no (k+1) collinear points, Discrete and Comp. Geom., (2013) 50, 811-820.

[SoTa] J. Solymosi and T. Tao, An incidence theorem in higher dimensions. Discrete Comput. Geom. 48 (2012), no. 2, 255-280

[SoTo] J. Solymosi and C. Toth, Distinct distances in the plane. The Micha Sharir birthday issue. Discrete Comput. Geom. 25 (2001), no. 4, 629-634.

[SST] J. Spencer, E. Szemerédi, and W. Trotter, Unit distances in the Euclidean plane. Graph theory and combinatorics (Cambridge, 1983), 293-303, Academic Press, London, 1984.

[St] E. Stein, Some problems in harmonic analysis. Harmonic analysis in Euclidean spaces (Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978), Part 1, pp. 3-20, Proc. Sympos. Pure Math., XXXV, Part, Amer. Math. Soc., Providence, R.I., 1979.

[StSh] E. Stein and R. Shakarchi, *Real analysis. Measure theory, integration, and Hilbert spaces*, Princeton Lectures in Analysis, III. Princeton University Press, Princeton, NJ, 2005.

[StTu] A. Stone and J. Tukey, Generalized "sandwich" theorems. Duke Math. J. 9, (1942). 356-359.

[Su] M. Sudan, Efficient checking of polynomials and proofs and the hardness of approximation problems, ACM Distinguished Thesees, Springer 1995.

[Sz] L. Székely, Crossing numbers and hard Erdős problems in discrete geometry. Combin. Probab. Comput. 6 (1997), no. 3, 353-358.

[SzTr] E. Szemerédi and W. T. Trotter Jr., Extremal Problems in Discrete Geometry, Combinatorica (1983) 3, 381-392.

[Ta1] T. Tao, From rotating needles to stability of waves: emerging connections between combinatorics, analysis, and PDE. Notices Amer. Math. Soc. 48 (2001), no. 3, 294-303.

[Ta2] T. Tao, Lecture notes on restriction, Math 254B, Spring 1999.

[To] C. Toth, The Szemerédi-Trotter theorem in the complex plane. aXiv:math/0305283, 2003.

[Tr] L. Trevisan, Some applications of coding theory in computational complexity. Complexity of computations and proofs, 347-424, Quad. Mat., 13, Dept. Math., Seconda Univ. Napoli, Caserta, 2004.

[Va] https://proofwiki.org/wiki/Vandermonde_Determinant

[Wo1] T. Wolff, An improved bound for Kakeya type maximal functions. Rev. Mat. Iberoamericana 11 (1995), no. 3, 651-674.

[Wo2] T. Wolff. Recent work connected with the Kakeya problem. Prospects in mathematics (Princeton, NJ, 1996). pages 129-162, 1999.

[Wo3] T. Wolff, *Lectures on Harmonic Analysis*, American Mathematical Society, University Lecture Series vol. 29, 2003.

[WYZ] H. Wang, B. Yang, and R. Zhang, Bounds of incidences between points and algebraic curves, arXiv:1308.0861

[Z] R. Zhang, On sharp local turns of planar polynomials. Math. Z. 277 (2014), no. 3-4, 1105-1112.

Selected Published Titles in This Series

64 **Larry Guth,** Polynomial Methods in Combinatorics, 2016

63 **Gonçalo Tabuada,** Noncommutative Motives, 2015

62 **H. Iwaniec,** Lectures on the Riemann Zeta Function, 2014

61 **Jacob P. Murre, Jan Nagel, and Chris A. M. Peters,** Lectures on the Theory of Pure Motives, 2013

60 **William H. Meeks III and Joaquín Pérez,** A Survey on Classical Minimal Surface Theory, 2012

59 **Sylvie Paycha,** Regularised Integrals, Sums and Traces, 2012

58 **Peter D. Lax and Lawrence Zalcman,** Complex Proofs of Real Theorems, 2012

57 **Frank Sottile,** Real Solutions to Equations from Geometry, 2011

56 **A. Ya. Helemskii,** Quantum Functional Analysis, 2010

55 **Oded Goldreich,** A Primer on Pseudorandom Generators, 2010

54 **John M. Mackay and Jeremy T. Tyson,** Conformal Dimension, 2010

53 **John W. Morgan and Frederick Tsz-Ho Fong,** Ricci Flow and Geometrization of 3-Manifolds, 2010

52 **Marian Aprodu and Jan Nagel,** Koszul Cohomology and Algebraic Geometry, 2010

51 **J. Ben Hough, Manjunath Krishnapur, Yuval Peres, and Bálint Virág,** Zeros of Gaussian Analytic Functions and Determinantal Point Processes, 2009

50 **John T. Baldwin,** Categoricity, 2009

49 **József Beck,** Inevitable Randomness in Discrete Mathematics, 2009

48 **Achill Schürmann,** Computational Geometry of Positive Definite Quadratic Forms, 2008

47 **Ernst Kunz, David A. Cox, and Alicia Dickenstein,** Residues and Duality for Projective Algebraic Varieties, 2008

46 **Lorenzo Sadun,** Topology of Tiling Spaces, 2008

45 **Matthew Baker, Brian Conrad, Samit Dasgupta, Kiran S. Kedlaya, and Jeremy Teitelbaum,** p-adic Geometry, 2008

44 **Vladimir Kanovei,** Borel Equivalence Relations, 2008

43 **Giuseppe Zampieri,** Complex Analysis and CR Geometry, 2008

42 **Holger Brenner, Jürgen Herzog, and Orlando Villamayor,** Three Lectures on Commutative Algebra, 2008

41 **James Haglund,** The q, t-Catalan Numbers and the Space of Diagonal Harmonics, 2008

40 **Vladimir Pestov,** Dynamics of Infinite-dimensional Groups, 2006

39 **Oscar Zariski,** The Moduli Problem for Plane Branches, 2006

38 **Lars V. Ahlfors,** Lectures on Quasiconformal Mappings, Second Edition, 2006

37 **Alexander Polishchuk and Leonid Positselski,** Quadratic Algebras, 2005

36 **Matilde Marcolli,** Arithmetic Noncommutative Geometry, 2005

35 **Luca Capogna, Carlos E. Kenig, and Loredana Lanzani,** Harmonic Measure, 2005

34 **E. B. Dynkin,** Superdiffusions and Positive Solutions of Nonlinear Partial Differential Equations, 2004

33 **Kristian Seip,** Interpolation and Sampling in Spaces of Analytic Functions, 2004

32 **Paul B. Larson,** The Stationary Tower, 2004

31 **John Roe,** Lectures on Coarse Geometry, 2003

30 **Anatole Katok,** Combinatorial Constructions in Ergodic Theory and Dynamics, 2003

29 **Thomas H. Wolff,** Lectures on Harmonic Analysis, 2003

28 **Skip Garibaldi, Alexander Merkurjev, and Jean-Pierre Serre,** Cohomological Invariants in Galois Cohomology, 2003

For a complete list of titles in this series, visit the
AMS Bookstore at **www.ams.org/bookstore/ulectseries/**.

This book explains some recent applications of the theory of polynomials and algebraic geometry to combinatorics and other areas of mathematics. One of the first results in this story is a short elegant solution of the Kakeya problem for finite fields, which was considered a deep and difficult problem in combinatorial geometry. The author also discusses in detail various problems in incidence geometry associated to Paul Erdős's famous distinct distances problem in the plane from the 1940s. The proof techniques are also connected to error-correcting codes, Fourier analysis, number theory, and differential geometry. Although the mathematics discussed in the book is deep and far-reaching, it should be accessible to first- and second-year graduate students and advanced undergraduates. The book contains approximately 100 exercises that further the reader's understanding of the main themes of the book.

Some of the greatest advances in geometric combinatorics and harmonic analysis in recent years have been accomplished using the polynomial method. Larry Guth gives a readable and timely exposition of this important topic, which is destined to influence a variety of critical developments in combinatorics, harmonic analysis and other areas for many years to come.

—Alex Iosevich, *University of Rochester, author of "The Erdős Distance Problem" and "A View from the Top"*

It is extremely challenging to present a current (and still very active) research area in a manner that a good mathematics undergraduate would be able to grasp after a reasonable effort, but the author is quite successful in this task, and this would be a book of value to both undergraduates and graduates.

—Terence Tao, *University of California, Los Angeles, author of "An Epsilon of Room I, II" and "Hilbert's Fifth Problem and Related Topics"*

ISBN 978-1-4704-2890-7

9 781470 428907

ULECT/64

For additional information and updates on this book, visit

www.ams.org/bookpages/ulect-64