The Erdős distance problem

Julia Garibaldi, Alex Iosevich,
and Steven Senger

Contents

Biographical information

The first listed author was born on October 2, 1976 in Seattle and was raised across the water on Bainbridge Island. She graduated from New York University in the Spring of 1999 and went on to UCLA to get her PhD in December 2004. She spent two years at Georgia Tech as a postdoctoral fellow and has had lecturing positions at Emory University since.

The second listed author was born in Lvov, USSR, on December 14, 1967, emigrated to the United States of America at the age of eleven with his immediate family, and grew up in Chicago, Illinois. He graduated from the University of Chicago in 1989 with a B.S. in Pure Mathematics, and a Ph.D. from UCLA in 1993 under the direction of Christopher Sogge. After appointments at McMaster University, Wright State University, and Georgetown, the author spent ten years at the University of Missouri, where this book was written. In July of 2010, he is moving to the University of Rochester.

The third listed author was born in North Kansas City, Missouri, on May 19, 1982. He graduated from the University of Missouri in 2005 with a B.S. in Computer Engineering, a B.S. in Electrical Engineering, and a B.S. in Mathematics. Bewteen various musical performances and rock climbing excursions, he is currently working on a Ph.D. in Mathematics under the direction of the second listed author.

Foreword

There are several goals for this book. As the title indicates, we certainly hope to familiarize you with the some of the major results in the study of the Erdős distance problem. This goal should be easily attainable for most experienced mathematicians. However, if you are not an experienced mathematician, we hope to guide you through many advanced mathematical concepts along the way. This book is based on the notes that were written for the summer program on the problem, held at the University of Missouri, August 1-5, 2005. This was the second year of the program, and our plan continued to be to introduce motivated high school students to accessible concepts of higher mathematics.

This book is designed to be enjoyed by readers at different levels of mathematical experience. Some of the notes and remarks are directed at graduate students and professionals in the field. So, if you are relatively inexperienced, and a particular comment or observation uses terminology[1] that you are not familiar with, you may want to skip past it or look up the definitions later. On the other hand, if you are a more experienced mathematician, feel free to skim the introductory portions to glean the necessary notation, and move on to the more specific subject matter.

[1] One example of this is the mention of curvature in the first section of the Introduction.

Our book is heavily problem oriented. Most of the learning is meant to be done by working through the exercises. Many of these exercises are recently published results by mathematicians working in the area. In a couple of places, steps are intentionally left out of proofs and, in the process of working the exercises, the reader is then asked to fill them in. On a number of occasions, solutions to exercises are used in the book in an essential way. Sometimes the exercises are left to the end of the chapter, but a few times, we intersperse them throughout the chapter to illustrate concepts or to get the reader's hands dirty, so the ideas really sink in right at that point in the exposition. Also, some exercises are much more complicated than others, and will probably require several hours of concentrated effort for even an advanced student. So please do not get discouraged. Having said that, let us add that you should not rely solely on exercises in these notes. Create your own problems and questions! Modify the lemmas and theorems below, and, whenever possible, improve them! Mathematics is a highly personal experience and you will find true fulfillment only when you make the concepts in these notes your own in some way. Read this book with a pad of paper handy to really explore these ideas as they come along. Good luck!

Introduction

Many theorems in mathematics say, in one way or another, that it is very difficult to arrange mathematical objects in such a way that they do not exhibit some interesting structure. The objects in the Erdős distance problem are points, and the structure we are curious about involves distances between points. We can loosely formulate the main question of this book as: How many distinct distances are determined by a finite set of points?

1. A sketch of our problem

In the case that there is only one point, we have but one distance, zero. In the case of two points, our job is pretty easy again. We have the distance between the two points, and again, zero. However, if we consider the case of three points in the plane, it starts to get interesting. Three points arranged as the vertices of an equilateral triangle are the same distance from one another, so there is only one non-zero distance, making two total. If they are the vertices of an isosceles triangle, we have one distance repeated, leaving three distinct distances total. Of course, there are any number of ways for three points to determine four distances. These phenomena increase in complexity and frequency as we consider more and more points. In fact, there is no configuration of four points that has only one nonzero distance present. It stands to reason that as we add more points, we will add more distances.

For the study of this problem, we fix a dimension to work in, d, and we then are concerned with the *asymptotic* behavior of n, or how things happen as n grows large, past a million, past a billion, and on. Since we are considering large n, we will not be concerned with the exact number of distinct distances, but about how many distinct distances there are in comparison to n.

To be precise, in full generality the Erdős distance problem asks for the minimal number of distances determined by a set, P, of n points in d dimensional space, $\mathbb{R}^d$. For this to be interesting, we will assume $d \geq 2$. Define

$$\Delta(P) := \{|p - p'| : p, p' \in P\},$$

and

$$|x| := \sqrt{x_1^2 + \cdots + x_d^2},$$

the standard Euclidean distance.

Using this notation, we want to know the smallest possible size of $\Delta(P)$ for some point set P. Let us consider some simple examples that involve many points. Let P consist solely of n points in a line along the x-axis. Then $\Delta(P) = \{0, 1, 2, \ldots, n-1\}$. This is because each point determines distances that are integers between 1 and n. This simple example shows that there is a set of n points that only determines about n distinct distances.

In general, we can construct sets of n points in d dimensions that only have about $n^{\frac{2}{d}}$ distinct distances. This will be outlined in Exercise 0.3. For two dimensions, it turns out that there can even be asymptotically fewer distinct distances than the size of P.

The conjecture is nowhere near resolution, but much is known, and we will come very close to the cutting edge of this beautiful problem in this book. One of the great things about this theory is that it can be developed largely from the ground up. That is, this problem in particular can be studied without a lot of background. So if you are curious as to what mathematical research is like, reading through this book can provide you a glimpse. You can actively watch the theory grow from its infancy through some of the most recent discoveries in the field. Along the way, you will be introduced to many of the elementary techniques in any serious mathematician's toolkit. If you are already familiar with research mathematics, and desire more justification for serious exploration of this particular area, we have included sketches of some consequences of the study of this problem in the final chapter of this book. More precisely, we show how the Erdős distance problem and the Erdős integer distance principle can be used to show that a set of mutually orthogonal exponentials on a smooth symmetric convex surface in $\mathbb{R}^d$ with everywhere non-vanishing curvature must be very small. This provides a connection between a set of problems in classical analysis and the main theme of this book. This is just one of many connections between the Erdős distance problems and other areas of mathematics. An interested reader is encouraged to consult a beautiful article by Nets Katz and Terry Tao ([**24**]) and the references contained therein.

2. Some notation

If you are not familiar with some of the mathematical notation used in this book, the following should serve as a quick reference.

As above, if x is a vector, $|x|$ will denote its length, or distance from the origin. Of course, if y is also a vector, $|x-y|$ will denote the distance between x and y.

We call a set of elements A. Denote the elements in the set as $A := \{a_1, a_2, ..., a_n\}$.[2] We can denote the size of the set as $|A|$, or sometimes $\#A$. Union and intersection are denoted as usual, with $\cup$ and $\cap$, respectively.

Suppose we have two sets, $A := \{2,4,6,8\}$, and $B := \{1,2,3,4,5,6\}$. Then $A \cup B = \{1,2,3,4,5,6,8\}$, and $A \cap B = \{2,4,6\}$. Also, we say that 1 is an element of B like this: $1 \in B$. Of course, 1 is not an element of A, so $1 \notin A$.

These operations can be indexed. Suppose that $A_1, A_2, ..., A_m$ denote a sequence of m sets. We can write an indexed union or intersection as follows:

[2] The colon next to the equals sign just means that this is a definition.

$$\bigcup_{i=1}^{m} A_i = A_1 \cup A_2 \cup ... \cup A_m$$

$$\bigcap_{i=1}^{m} A_i = A_1 \cap A_2 \cap ... \cap A_m$$

Similarly, if we have a sequence of numbers, $a_1, a_2, ..., a_m$ be a sequence of m numbers. We can compute their indexed sum as follows:

$$\sum_{i=1}^{m} a_i = a_1 + a_2 + ... + a_m.$$

If the context is clear, this may be abbreviated as

$$\sum_i a_i.$$

We will denote the binomial coefficient as $\binom{n}{k}$, and it means

$$\frac{n!}{k!(n-k)!},$$

which is the number of ways to choose k objects from n.

Here, and throughout the book, $X \lesssim Y$ means that as X and Y grow large, typically as a function of some parameter, there exists a positive constant C such that $X \leq CY$, and $X \approx Y$ means that $X \lesssim Y$ and $Y \lesssim X$. We take this notational game a step further and define $X \lessapprox Y$, with respect to the large parameter N, if for every $\epsilon > 0$ there exists $C_\epsilon > 0$ such that $X \leq C_\epsilon N^\epsilon Y$. This notation is not only more convenient, but it also emphasizes that many of these constants do not affect our results asymptotically.

Naturally, as the the theory develops, we will use more symbols and shorthand, but these will all be introduced as they arise.

Now we state the Erdős distance conjecture formally, with the notation used in this book.

Erdős distance conjecture Let P be a subset of $\mathbb{R}^d$, $d \geq 2$, such that $\#P = n$. Then

$$\#\Delta(P) \gtrapprox n, \text{ if } d = 2, \tag{0.1}$$

and

$$\#\Delta(P) \gtrsim n^{\frac{2}{d}}, \text{ if } d \geq 3. \tag{0.2}$$

Exercises

Exercise 0.1. Suppose there are p pigeons, each huddled in one of h holes, with $p > h$. Explain why there must be at least one hole with at least $\frac{p}{h}$ pigeons in it. This is known as the *pigeonhole principle.*

Exercise 0.2. Determine the minimum number of distances determined by n points in the plane for $n = 3, 4$ and 5. How do things change for points in three-dimensional space?

Exercise 0.3. Let $P = \mathbb{Z}^d \cap [0, n^{\frac{1}{d}}]^d$, where n is a dth power of an integer. Then $\Delta(P) = \{|p| : p \in P\}$ (why?) and $\#\Delta(P) = \#\{|p|^2 : p \in P\}$. Consider the set of numbers $p_1^2 + p_2^2 + \cdots + p_d^2$, $p = (p_1, \ldots, p_d) \in P$. All these numbers are positive integers no less than 0 and no more than $dn^{\frac{2}{d}}$. Now check that

$$\#\Delta(P) \leq dn^{\frac{2}{d}} + 1$$

follows from this observation.

Exercise 0.4. Define $\Delta_{l_1(\mathbb{R}^d)}(P) = \{|p_1 - p_1'| + \cdots + |p_d - p_d'| : p, p' \in P\}$. Prove that Erdős distance conjecture is false if $\Delta(P)$ is replaced by $\Delta_{l_1(\mathbb{R}^d)}(P)$. What should the conjecture say in this context? Consider the case $d = 2$ first.

Exercise 0.5. Let K be a *convex*, *centrally symmetric* subset of $\mathbb{R}^2$, contained in the disk of radius 2 centered at the origin and containing the disk of radius 1 centered at the origin. Convex means that if x and y are points in K, then the line segment connecting x and y is contained entirely inside K. Centrally symmetric means that if x is in K, then $-x$ is also in K.

Let $t = ||x||_K$ denote the number such that x is contained in tK, but is not contained in $(t - \epsilon)K$ for any $\epsilon > 0$. Define $\Delta_K(P) = \{||p - p'||_K : p, p' \in P\}$. If the boundary of K contains a line segment, prove that one can construct a set, P, with $\#P = n$, such that $\#\Delta_K(P) \lesssim n^{\frac{1}{d}}$. This is called the *Minkowski functional.*

Chapter 1

The $\sqrt{n}$ theory

1. Erdős' original argument

How does one prove that any set, P, of size n determines many distances? Let us start in two dimensions. We will begin by giving two proofs of the following theorem– the first of which was originally published by Erdős in 1946.

Theorem 1.1 (Erdős [**9**]). *Suppose that $d = 2$ and $\#P = n$. Then $\#\Delta(P) \gtrsim n^{\frac{1}{2}}$.*

1st proof. Chose a point, p_0, and draw circles around it that each contain at least one point of P. Continue drawing circles around p_0 until all the points in P lie on a circle of some radius centered at p_0. We will refer to this procdure as *covering* the points of P by circles centered at p_0. We can think of each circle as a *level set*, or set of points that have the same value for some funciton. In this case, the function is the distance from the point p_0. Suppose that we have drawn t circles. This means that we can be assured that there are at least t different distances between points in P and p_0. If t is greater than $n^{\frac{1}{2}}$, then we are already doing very well. But what if t is happens to be small? Note that at least one of the t circles must

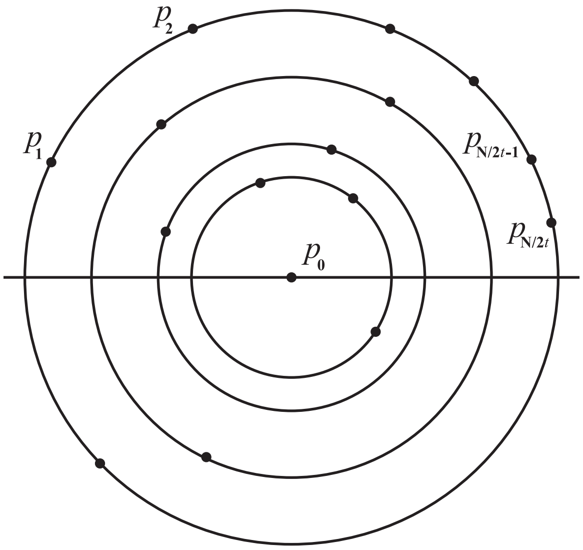


Figure 1.1. Circles about p_0 and the East-West line.

contain at least n/t points[1] , by the pigeonhole principle. Draw the East-West line though the center of that circle. Then at least $n/2t$ are contained in either the Northern or Southern hemisphere. Without loss of generality[2], suppose that there are $n/2t$ points in the Northern hemisphere.

Fix the East-most point and draw segments from that point to all the other points of P in the Northern hemisphere. The lengths of these segments are all different, so at least $n/2t$ distances are thus

[1] Actually, this would be $\frac{n-1}{t}$ points, but since $\frac{n-1}{t} \approx \frac{n}{t}$, we will continue with the simpler notation. This may seem annoying, but it is done intentionally to keep the most important information at the forefront.

[2] As in many proofs, we are asserting something "without loss of generality", which is often abbreviated WLOG. What this typically means is that we can simplify the notation of the proof to get to the point, and we let the reader fill in the trivial details later. In this instance, it means that we can deal with the case that most of the points are in the Northern hemisphere. If they were in the Southern hemisphere, the proof would not change much, we would just restate it, word for word, but say Southern instead of Northern from this point onward.

determined. This proves that

$$\#\Delta(P) \geq \max\{t, n/2t\}. \tag{1.1}$$

There are several ways to proceed here. One way is to "guess" the answer. Since we already took care of the case where $t \geq \sqrt{n}$, we can assume that $t < \sqrt{n}$. Then $n/2t > \sqrt{n}/2$, so either way,

$$\#\Delta(P) \gtrsim \sqrt{n}. \tag{1.2}$$

A slightly less "sneaky" approach is to use the fact that

$$\max\{X, Y\} \geq \sqrt{XY} \text{ (why?)}.$$

This transforms (1.1) into (1.2). □

2nd proof. Take any two points p_1 and p_2 from P. Draw in the circles about p_1 and p_2 that capture the remaining $N-2$ points of P. Suppose there are t circles about p_1 and s circles about p_2. Since all of the points of P are in the intersections of these two families of circles, we have that $n-2 \leq 2st$ (why?). Therefore either $s \gtrsim \sqrt{n}$ or $t \gtrsim \sqrt{n}$ and we are done. □

2. Higher dimensions

What about higher dimensions? Let us try the same approach. Choose a point in P and draw all spheres that contain at least one point of P. As before, let t denote the number of spheres. If t is large enough, we are done. If not, then one of the spheres contains at least n/t points. Unfortunately, if $d > 2$, we cannot run the simple minded argument that worked in two dimensions. Or can we? Notice that if we are working in $\mathbb{R}^d$, the surface of each sphere is $(d-1)$-dimensional, whatever that means. This suggests the following approach, which uses induction. If you are unfamiliar with proofs by induction, Appendix C has an brief explanation of this concept.

Proposition 1.2 (Induction Hypothesis). *Let P' be a subset of $\mathbb{R}^k$, $k \geq 2$, or P' is a subset of S^k, $k \geq 1$. Suppose that $\#P' = n'$. Then*

$$\#\Delta(P') \gtrsim (n')^{\frac{1}{k}}.$$

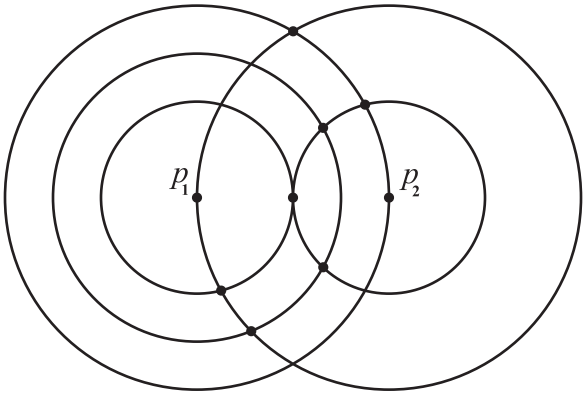


Figure 1.2. Circles about p_1 and p_2 that cover P.

In the case of $\mathbb{R}^k$, the induction hypothesis holds if $k = 2$, as we have verified above. Similarly, we have verified the statement for S^k for $k = 1$. We are now ready to complete the higher dimensional argument. When we follow this reasoning in dimension d, we end up with t $(d-1)$-spheres–one of which must have at least n/t points on it as in the $d = 2$ proof. By induction, these points determine $\gtrsim \left(\frac{n}{t}\right)^{\frac{1}{d-1}}$ distances. It follows that

$$\#\Delta(P) \gtrsim \max\left\{t, \left(\frac{n}{t}\right)^{\frac{1}{d-1}}\right\}.$$

We now use the fact that

$$\max\{X, Y\} \geq (XY^{d-1})^{\frac{1}{d}} \text{ (why?)},$$

which implies that

$$\#\Delta(P) \gtrsim n^{\frac{1}{d}}. \tag{1.3}$$

We just proved the following result.

Theorem 1.3. *Let P be a subset of $\mathbb{R}^d$, $d \geq 2$, such that $\#P = n$. Then $\#\Delta(P) \gtrsim n^{\frac{1}{d}}$.*

3. Arbitrary metrics

Although we have been mostly thinking about the standard Euclidean metric so far, it is possible to consider other metrics. For example, what if you were walking from the corner of one city block to the corner of another, say a street corner three blocks north and four blocks east? It is most likely that you could not just take a direct route along the straight line connecting the two corners. There are probably buildings in the way. You would probably do something like walk north for two blocks, and then walk east for two blocks. Even though the "distance" between the two street corners seemed to be about five blocks, you end up walking seven blocks. This is one way of thinking about the l_1 metric mentioned in Exerceise 0.4. It is sometimes referred to as the *taxicab* or *Manhattan* metric.

We now present a formal definition of a general metric.

Definition 1.4. We call a function, $d(x, y)$, on a set, S, a *metric* if it returns a real number for any two elements of S satisfying the following for all distinct $x, y, z \in S$:
(i) $d(x, x) = 0$
(ii) $d(x, y) > 0$
(iii) $d(x, y) = d(y, x)$ (symmetry)
(iv) $d(x, z) \geq d(x, y) + d(y, z)$ (Triangle Inequality)

Dropping the symmetry assumption from the definition gives us a similar object called an *asymmetric metric.* Many of the arguments to follow do not depend heavily on the symmetry of the metric. When you are comfortable with the general ideas in this book, see how many can still yield non-trivial results with asymmetric metrics.

We will explore this further in Chapter 5, but until then, just use your imagination as to what kinds of restrictions we will need for the proof ideas to go through.

It is customary to think of the distance from one point to another as the length of the straight line connecting the two points. However, as our cursory exploration of the taxicab metric suggests, this does not shed much light on how different metrics behave with respect to one another. One way to get a feel for a metric's behavior is by looking at its "spheres". If you fix one point, x, and consider the *locus*, or graphical representation, of points that are a given distance from x, using the standard Euclidean distance, you will get a sphere. Of course, a sphere in the plane is a circle. What would a such a "circle" look like in the l_1 metric? As you can see in Figure 1.3, the circles look like diamonds, or squares that have been rotated 45 degrees.

Now, this all depends on the circles or spheres of each respective metric looking the same throughout the space they are drawn in. For example, if you were to measure the length of a stick in El Paso, and then measure the length of the same stick in Chicago, you would expect the length to be the same. This property is called *homogeneity*.

In the arguments above, not all of the properties of the standard Euclidean circle were utilized. Exercises 1.6 and 1.7 will accentuate some of the critical similarities and differences between arbitrary metrics and the Euclidean metric.

At this point, we could spend a long time introducing and developing many different types of metrics, but instead, we want you to discover on your own what types of objects can be viewed as metrics, and in what sense. As you read through this book, other types of metrics and metric-like objects will naturally come along. In mathematics, we don't often have some arbitrary definition of an object, and then explore it. Typically, various scenarios give rise to sensible constraints on a useful object, which are then compiled into a definition after the subject has been explored a little. For this book in particular, we feel that it is far more instructive to watch the theory grow by necessity than to introduce a laundry list of definitions and then draw conclusions. If you can come up with some of your own variations on the examples given in Exercise 1.8, you will get more out of this book.

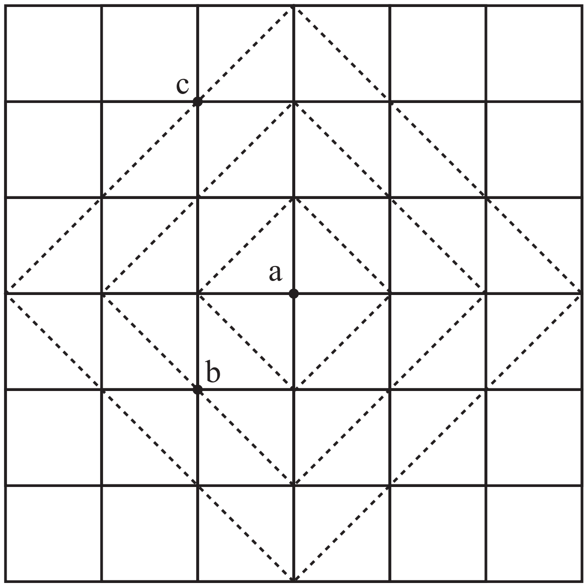


Figure 1.3. The grid represents an overhead view of a city. If you are located at a, you will have walk two blocks to b, or three blocks to c. The dotted lines represent three dilates of the l_1 circle.

Exercises

Exercise 1.1. Prove that the minimum of $\max\{t, n/2t\}$ is in fact $\sqrt{n}$. In other words, show that Erdős' method of proof cannot do better than $\#\Delta(P) \gtrsim \sqrt{n}$.

Exercise 1.2. Calculate the constants from the two different proofs of Theorem 1.1. In other words, find the smallest constant C is each proof such that $\#\Delta(P) \geq C\sqrt{n}$. Which proof gives a stronger result?

Exercise 1.3. Attempt to extend Theorem 1.1 to the l_1 metric defined in Exercise 0.4. Do either of the proofs work verbatim for this metric? If not, can either of the proofs be modified to obtain a result?

Exercise 1.4. We outline an alternate proof of Theorem 1.1. Let M_n denote the matrix constructed as follows. Fix $t \in \Delta(P)$ and let the entry $a_{pp'} = 1$ if $|p - p'| = t$, and 0 otherwise. Observe that for a fixed pair (p', p''), $p' \neq p''$, $a_{pp'} \cdot a_{pp''} = 1$ for at most one value of p (why?). Use this along with the Cauchy-Schwarz inequality (Detailed in Chapter 3.) to prove that $\sum_{p,p' \in P} a_{pp'} \lesssim n^{\frac{3}{2}}$. Conclude that for any $t \in \Delta(P)$, $\#\{(p,p') : |p - p'| = t\} \lesssim n^{\frac{3}{2}}$. Deduce that $\#\Delta(P) \gtrsim \sqrt{n}$. Can you make this idea run in higher dimensions?

Exercise 1.5. In the proofs of Theorems 1.1 and 1.3, we only used spheres centered at a single point. Is there any milage to be gained by considering, in some way, two points? Try it.

Exercise 1.6. Let K be a polygon in the plane. Let $\#P = n$. Let $\Delta_K(P) = \{||p - p'||_K : p, p' \in P\}$. Prove that $\#\Delta_K(P) \gtrsim \sqrt{n}$. What about other convex K?

Exercise 1.7. Why do the K in Exercise 1.6 have to be convex?

Exercise 1.8. Consider the following metric-like objects. Assume that they all map $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, or that they take two points in the plane as input and give one number as output. Determine which are genuine metrics, and which are not. Could one sensibly ask questions like the Erdős distance problem of these objects?

$$\begin{aligned} If x = (x_1, x_2), and y = (y_1, y_2), i)F(x,y) &= |x| + |y| \\ ii)D(x,y) &= x_1 x_2 + y_1 y_2 \\ iii)\Phi(x,y) &= \frac{|x - y|}{|x + y|} \end{aligned}$$

The first one is sometimes referred to as the *French Railroad* metric. The second is the standard dot product of x and y.

Exercise 1.9. Consider x, y, and $z \in \mathbb{R}^n$. Suppose $x \neq y$. If there is a function, $d : \mathbb{R}^n \to \mathbb{R}$, where $d(x,y) \neq d(x - z, y - z)$, can d be a metric? In this example, d could be described as *inhomogeneous*.

Chapter 2

The $n^{2/3}$ theory

1. The Erdős integer distance principle

Erdős' ingenious argument, described in the previous chapter, relies on spheres centered at a single point, and it stands to reason that one might gain something out of considering spheres centered at two points. This point of view was introduced by Leo Moser in the early 1950s. Before presenting Moser's argument, we will present the Erdős Integer Distance Principle, where an idea similar to Moser's is already present, albeit in a different form and context.

Erdős integer distance principle, [**10**]**.** Let A be an infinite subset of $\mathbb{R}^d$, $d \geq 2$. Suppose that $\Delta(A) \subset \mathbb{Z}$. Then A is contained in a line. □

We will prove this result by way of contradiction, which is sometimes abbreviated, "BWOC". This means that we will begin by assuming that our assertion is false, and use this to reason our way into a contradiction, or a situation that cannot be true. Since the assumption that our assertion was false yields faulty results, we conclude that our assertion must have been true after all.

To prove Erdős Integer Distance Principle, consider the possibility that A is not contained in a line. Suppose that $d = 2$. Let a, a', a'' denote three points of A that are not *collinear*, or not lying on the same line. Let b be any other point of A. By assumption, $|a-b|$ and $|a'-b|$ are both integers, which means that $|a-b|-|a'-b|$ is also an integer. This means that there is a collection hyperbolas with focal points at a and a', such that each point in A is on a hyperbola in the collection. (See Appendix A for a thorough description of basic theory of hyperbolas in the plane). How many such hyperbolas are there? Well, suppose that $|a-a'|=k$, which, by assumption is an integer. By the triangle inequality, $||a-b|-|a'-b|| \leq |a-a'|=k$. It follows that there are only $k+1$ different hyperbolas with focal points at a and a'. Similarly, all the points of A are contained in $l+1$ hyperbolas with focal points at a' and a''. Any hyperbola with focal points at a and a' and a hyperbola with focal points at a' and a'' intersect at at most 4 points (see the Exercise in Appendix A). If we let l be $|a'-a''|$, it follows that the number of points in A cannot exceed $16(k+1)(l+1)$, which is a contradiction since A is assumed to be infinite. This proves the two-dimensional case of the Erdős integer distance principle. The higher dimensional argument is outlined in Exercise 2.5 below.

The following beautiful extension of the Erdős integer distance principle was proved by Jozsef Solymosi [**39**].

Theorem 2.1. *Suppose that P is a subset of $\mathbb{R}^2$, such that $\Delta(P) \subset \mathbb{Z}$ and $\#P = n$. Suppose that P is contained in a disk of radius R. Then $R \gtrsim n$.*

The proof of Solymosi's theorem is outlined in Exercise 2.3, and in Exercise 2.4 we ask you to verify that Theorem 2.1 would follow immediately from the Erdős distance conjecture.

2. Moser's construction

We are now ready to introduce Moser's idea. You will probably notice that this proof is intentionally written in a highly symbolic, set-notational style. There are several reasons for this. It is important to see how little this argument has to do with many of the specific geometric qualities of circles. Since it is written so abstractly, it should be easier to pick out the key features of the geometry that are necessary for such an argument, so you can generalize it on your own.

Exercise 2.8 is one way to explore that. Also, the sooner you learn to cope with multiple definitions and indices flying around, the better. Math is not read left to right, top to bottom. You will probably have to re-read portions of this argument again and again until it all sinks in. Finally, this particular approach will set the reader up nicely for the types of ideas employed in the next Chapter.

Choose points X and Y in P such that

$$|X - Y| \leq \min\{|p - p'| : p, p' \in P\}.$$

Let O be the midpoint of the segment XY. Half of the points of P are either above or below the line connecting X and Y. Call this set of points P'. Assume without loss of generality that at least half the points are above the line. Draw annuli centered at O of thickness $|X - Y|$ until all the points of P' are covered.

Keep only one third of the annuli in such a way that at least one third of the points of P' are contained, and such that if a particular annulus is kept, the next two consecutive annuli are discarded. (Prove that this can be done and try to figure out why we are doing this as you read the rest of the argument!). Call the resulting set of points P''.

Our next step will be to consider what happens inside of each of the annuli that we kept. Notice that distances from X and Y to points in one annulus cannot occur in another of the annuli we kept. So we can count the distinct distances that we find in each annulus, since they can not be present in any of the other annuli we consider.

Let $\mathcal{A}_j$ denote the points of P'' in the jth annulus. Call the number of points in the jth annumlus n_j. If every point gave different distances, to X and Y, we would be quite happy. Since that might not be the case, suppose that there are k numbers such that

$$\{|p - X| : p \in \mathcal{A}_j\} \cup \{|p - Y| : p \in \mathcal{A}_j\} = \{d_1, d_2, \ldots, d_k\}.$$

So we are counting k distinct distances to both X and Y from points in the jth annulus. For now, let us concern ourselves only with the jth annulus. We will count the number of distinct distances in it, and then sum over all annuli later.

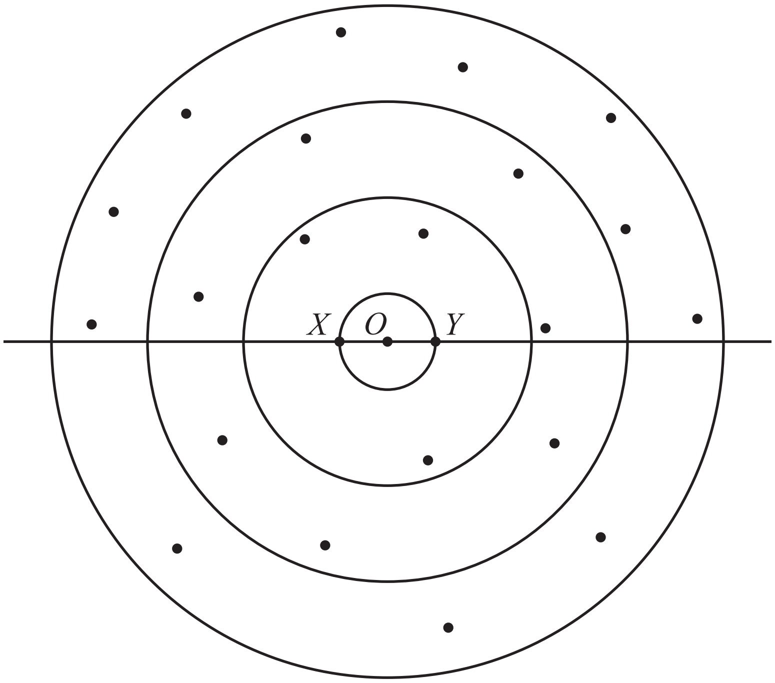


Figure 2.1. Annuli centered at O, the midpoint of X and Y of thickness $|X - Y|$.

Let

$$A_l = \{p \in \mathcal{A}_j : |p - X| = d_l\},$$

and

$$B_i = \{p \in \mathcal{A}_j : |p - Y| = d_i\}.$$

These are the sets of points in the jth annulus that lie on a circle of a given radius from X or Y.

By construction,

$$A_l = \bigcup_i (A_l \cap B_i),$$

since points of distance d_l from X are of some distance or another from Y. If we look at all distances to X in the jth annulus, it follows that

$$\bigcup_l A_l = \bigcup_{i,l} (A_l \cap B_i) \, .$$

Now,

$$\# \bigcup_l A_l = n_j ,$$

while

$$\# \bigcup_{i,l} (A_l \cap B_i) \leq k^2 \max_{i,l} \# (A_l \cap B_i) \, . \tag{2.1}$$

Now, A_l and B_i are contained on circles of approximately the same radius centered at different points, so $\max_{i,l} \#(A_l \cap B_i) \leq 1$. Plugging this into Equation 2.1, we see that

$$k \geq \sqrt{n_j} ,$$

from which we deduce that

$$\#\Delta(P) \geq \#\Delta(P'') \geq \sum_j \sqrt{n_j} . \tag{2.2}$$

We have

$$\frac{n}{6} \leq \sum_j n_j = \sum_j \sqrt{n_j} \cdot \sqrt{n_j} \leq \sqrt{n_{max}} \cdot \sum_j \sqrt{n_j} ,$$

where

$$n_{max} = \max_j n_j ,$$

which is the largest value of all of the n_j's. Observe that by the proof of Theorem 1.1,

$$\#\Delta(P) \geq \#\Delta(P'') \geq n_{max} .$$

By (2.2),

$$\#\Delta(P) \geq \frac{n}{6\sqrt{n_{max}}} .$$

It follows that

$$(\#\Delta(P))^2 \cdot \#\Delta(P) \geq n_{max} \cdot \frac{n^2}{36n_{max}} = \frac{n^2}{36}.$$

Which implies that

$$\#\Delta(P) \geq \frac{n^{\frac{2}{3}}}{(36)^{\frac{1}{3}}},$$

and we have just proved the following theorem.

Theorem 2.2 (Moser [**33**]). *Let $d = 2$ and suppose that $\#P = n$. Then $\#\Delta(P) \gtrsim n^{\frac{2}{3}}$.*

Exercises

Exercise 2.1. Outline of proof of Erdős Integer Distance Principle in higher dimensions.

Exercise 2.2. Prove that for every set of n points in the plane with diameter Δ and with at most $n/2$ collinear points, there exist two pairs of points A,B and C,D such that each of the distances $\overline{AB}$ and $\overline{CD}$ are less than $6\Delta/n^{1/2}$. *Hint:* Show that there are fewer than $n/2$ points that are not within $6\Delta/n^{1/2}$ of other points.

Exercise 2.3. Deduce Solymosi's Theorem from the the previous exercise by using ideas from the proof of the Erdős Integer Distance Principle.

Exercise 2.4. Deduce Solymosi's Theorem from the Erdős distance conjecture.

Exercise 2.5. Why did we eliminate 2/3 of the annuli in the proof above? Where did we use this in the proof?

Exercise 2.6. What does Moser's method yield in higher dimensions? Can you use the two-dimensional result along with the induction argument used to prove Theorem 1.3 instead? Which approach yields better exponents?

Exercise 2.7. Let A be an infinite subset of $\mathbb{R}^d$, $d \geq 2$, with the following property. We assume that $|a - a'| \geq \frac{1}{100}$ for all $a \neq a' \in A$. We also assume that for every $m \in \mathbb{Z}^d$, $[0,1]^d + m$ contains exactly one point of A. Let $A_q = [0,q]^d \cap A$. What kind of a bound can you obtain for $\Delta(A_q)$ using Moser's idea? Why is this bound better than the one we obtain above?

Take this a step further. Instead of using two points as in Moser's argument, use d points. How should these points be arranged? What effect are we trying to achieve? Can you obtain a better exponent this way?

Exercise 2.8. What happens if you try Moser's construction on the l_1 metric? What crucial difference keeps it from yielding greater exponents with this plan of attack? Can you imagine some reasonable conditions on metrics such that they would gain in Moser's construction over $n^{\frac{1}{2}}$?

Chapter 3

The Cauchy-Schwarz inequality

1. Proof of Cauchy-Schwarz

In this section we shall follow a procedure often considered nasty, but the one we hope to convince you to appreciate. We shall work backwards, discovering concepts as we go along, instead of stating them ahead of time. Let a and b denote two real numbers. Then

$$(a-b)^2 \geq 0. \tag{3.1}$$

This statement is so vacuous, you are probably wondering why we are telling you this. Nevertheless, expland the left hand side of (3.1). We get

$$a^2 - 2ab + b^2 \geq 0,$$

which implies that

$$ab \leq \frac{a^2+b^2}{2}. \tag{3.2}$$

Now consider two sums,

$$A_n = \sum_{k=1}^{n} a_k = a_1 + \cdots + a_n, \; B_n = \sum_{k=1}^{n} b_k = b_1 + \cdots + b_n,$$

where $a_1, \ldots, a_n$, and $b_1, \ldots, b_n$ are real numbers. Let

$$X_n = \left(\sum_{k=1}^{n} a_k^2\right)^{1/2} \quad \text{and} \quad Y_n = \left(\sum_{k=1}^{n} b_k^2\right)^{1/2}.$$

Our goal is to take advantage of (3.2). Let's take a look at

$$\begin{aligned} \sum_{k=1}^{n} a_k b_k &= X_n Y_n \sum_{k=1}^{n} \frac{a_k}{X_n} \cdot \frac{b_k}{Y_n} \\ &\leq X_n Y_n \sum_{k=1}^{n} \left[\frac{1}{2}\left(\frac{a_k}{X_n}\right)^2 + \frac{1}{2}\left(\frac{b_k}{Y_n}\right)^2\right]. \end{aligned} \tag{3.3}$$

Exercise 3.1. Explain using complete English sentences how (3.2) follows from 3.3.

Exercise 3.2. Explain why if C is a constant, then $\sum_{k=1}^{n} C a_k = C \sum_{k=1}^{n} a_k$.

Exercise 3.3. Explain why $\sum_{k=1}^{n} (a_k + b_k) = \sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k$.

We now use (3.2) and (3.3) to rewrite (3.3) in the form

$$\begin{aligned} X_n Y_n \frac{1}{2} \frac{1}{X_n^2} &\sum_{k=1} a_k^2 + X_n Y_n \frac{1}{2} \frac{1}{Y_n^2} \sum_{k=1}^{n} b_k^2 \\ &= X_n Y_n \frac{1}{2} \frac{1}{X_n^2} X_n^2 + X_n Y_n \frac{1}{2} \frac{1}{Y_n^2} Y_n^2 \\ &= \frac{1}{2} X_n Y_n + \frac{1}{2} X_n Y_n = X_n Y_n. \end{aligned}$$

Putting everything together, we have shown that

$$\sum_{k=1}^{n} a_k b_k \leq \left(\sum_{k=1}^{n} a_k^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^{n} b_k^2\right)^{\frac{1}{2}}. \tag{3.4}$$

This is known as the *Cauchy-Schwartz Inequality.*

Exercise 3.4. (This exercise is quite difficult if you do not know calculus) Let $1 < p < \infty$ and define the exponent p' by the equation $\frac{1}{p} + \frac{1}{p'} = 1$. Then

$$\sum_{k=1}^{n} a_k b_k \leq \left(\sum_{k=1}^{n} |a_k|^p\right)^{1/p} \left(\sum_{k=1}^{n} |b_k|^{p'}\right)^{1/p'}. \tag{3.5}$$

Observe that (3.5) reduces to (3.4) if $p = 2$. *Hint*: prove that $ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'}$ and proceed as in the case $p = 2$. One way to prove this inequality is to set $a^p = e^x$ and $b^{p'} = e^y$ (why are we allowed to do that?). Let $\frac{1}{p} = t$ and observe that $0 \leq t \leq 1$. We are then reduced to showing that for any real valued x, y and $t \in [0, 1]$, $e^{tx+(1-t)y} \leq te^x + (1-t)e^y$. This is exactly what it means for a function to be *convex*. Let $f(t) = e^{tx+(1-t)y}$ and $g(t) = te^x + (1-t)e^y$. Observe that $f(0) = g(0) = e^y$ and $f(1) = g(1) = e^x$. Can you complete the argument?

2. Application: Projections

Let's now try to see what the Cauchy-Schwartz (C-S) Inequaity is good for. Let S_n be a finite set of n points in $\mathbb{R}^3 = \{(x_1, x_2, x_3) : x_j \text{ is a real number}\}$, the three-dimensional Euclidean space. Let $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ and define

$$\pi_1(x) = (x_2, x_3),\ \pi_2(x) = (x_1, x_3),\ \text{and } \pi_3(x) = (x_1, x_2).$$

These are called *projections*. If we consider a point p in three dimensions, then $\pi_1(p)$ is like the "shadow" of p on the "wall" represented by the yz-plane. The question we ask is the following. We are assuming that $\#S_n = n$. What can we say about the size of $\pi_1(S_n), \pi_2(S_n)$, and $\pi_3(S_n)$? Before we do anything remotely complicated, let's make up some silly looking examples and see what we can learn from them.

Let $S_n = \{(0, 0, k) : k \text{ integer } k = 0, 1, \ldots, n-1\}$. This set clearly has n elements. What is $\pi_3(S_n)$ in this case. It is precisely the set $\{(0, 0)\}$, a set consisting of one element. However, $\pi_2(S_n)$ and $\pi_1(S_n)$ are both $\{(0, k) : k = 0, 1, \ldots, n-1\}$, sets consisting of n elements. In summary, one of the projections is really small and the others are as large as they can be.

Let's be a bit more even handed. Let $S_n = \{(k, l, 0) : k, l \text{ integers } 1 \leq k \leq \sqrt{n}, 1 \leq l \leq \sqrt{n}\}$, where $\sqrt{n}$ is an integer. Again $\#S_n = n$. What do projections look like? Well, S_n is already in the (x_1, x_2)-plane, so $\pi_3(S_n) = \{(k, l) : k, l \text{ integers } 1 \leq k \leq \sqrt{n}, 1 \leq l \leq \sqrt{n}\}$. It follows that $\#\pi_3(S_n) = n$. On the other hand, $\pi_2(S_n) = \{(k, 0) :$

k integer $1 \le k \le \sqrt{n}\}$, and $\pi_1(S_n) = \{(l, 0) : l$ integer $1 \le l \le \sqrt{n}\}$, both containing $\sqrt{n}$ elements. Again we see that it is difficult for all of the projections to be small.

Let's think about our examples so far from a geometric point of view. The first example is "one-dimensional" since the points all lie on a line. The second example is "two-dimensional" since the points lie on a plane. Let's now build a truly "three-dimensional" example with as much symmetry as possible. Let $S_n = \{(k, l, m) : k, l, m$ integers $1 \le k, l, m \le n^{\frac{1}{3}}\}$, where $n^{\frac{1}{3}}$ is an integer. Again, $\#S_n = n$, as required. This time the projections all look the same. We have $\pi_1(S_n) = \{(l, m) : l, m$ integers $1 \le l, m \le n^{\frac{1}{3}}\}$, a set of size $n^{\frac{2}{3}}$, and the same is true of $\#\pi_2(S_n)$ and $\#\pi_3(S_n)$.

Let's summarize what happened. In the case when all the projections have the same size, each projection has $n^{\frac{2}{3}}$ elements. We will see in a moment that for any S_n, one of the projections must of size at least $n^{\frac{2}{3}}$. Here and later in this book, we will see that the Cauchy-Schwarz inequality is very usefull in showing that the "symmetric" case is "optimal", whatever that means in a given instance.

To start our investigation, we need the following basic definition. Let S be any set. Define $\chi_S(x) = 1$ if $x \in S$ and 0 otherwise.

Exercise 3.5. Let S_n be as above. Then

$$\chi_{S_n}(x) \le \chi_{\pi_1(S_n)}(x_2, x_3)\chi_{\pi_2(S_n)}(x_1, x_3)\chi_{\pi_3(S_n)}(x_1, x_2).$$

With exercise 3.5 in tow, we write

$$n = \#S_n = \sum_x \chi_{S_n}(x) \le \sum_x \chi_{\pi_1(S_n)}(x_2, x_3)\chi_{\pi_2(S_n)}(x_1, x_3)\chi_{\pi_3(S_n)}(x_1, x_2)$$

$$= \sum_{x_1, x_2} \chi_{\pi_3(S_n)}(x_1, x_2) \sum_{x_3} \chi_{\pi_1(S_n)}(x_2, x_3)\chi_{\pi_2(S_n)}(x_1, x_3)$$

$$\le \left(\sum_{x_1, x_2} \chi^2_{\pi_3(S_n)}(x_1, x_2) \right)^{\frac{1}{2}} \left(\sum_{x_1, x_2} \left(\sum_{x_3} \chi_{\pi_1(S_n)}(x_2, x_3)\chi_{\pi_2(S_n)}(x_1, x_3) \right)^2 \right)^{\frac{1}{2}}$$

$$= I \times II.$$

Now,

$$I = \left(\sum_{x_1,x_2} \chi^2_{\pi_3(S_n)}(x_1,x_2)\right)^{\frac{1}{2}} = \left(\sum_{x_1,x_2} \chi_{\pi_3(S_n)}(x_1,x_2)\right)^{\frac{1}{2}} = (\#\pi_3(S_n))^{\frac{1}{2}}.$$

On the other hand,

$$II^2 = \sum_{x_1,x_2} \left(\sum_{x_3} \chi_{\pi_1(S_n)}(x_2,x_3)\chi_{\pi_2(S_n)}(x_1,x_3)\right)^2$$

$$= \sum_{x_1,x_2}\sum_{x_3}\sum_{x_3'} \chi_{\pi_1(S_n)}(x_2,x_3)\chi_{\pi_2(S_n)}(x_1,x_3)\chi_{\pi_1(S_n)}(x_2,x_3')\chi_{\pi_2(S_n)}(x_1,x_3')$$

$$\leq \sum_{x_1,x_2}\sum_{x_3}\sum_{x_3'} \chi_{\pi_1(S_n)}(x_2,x_3)\chi_{\pi_2(S_n)}(x_1,x_3')$$

$$= \sum_{x_2,x_3} \chi_{\pi_1(S_n)}(x_2,x_3) \sum_{x_1,x_3'} \chi_{\pi_2(S_n)}(x_1,x_3') = \#\pi_1(S_n)\cdot\#\pi_2(S_n).$$

Putting everything together, we have proved that

$$\#S_n \leq \sqrt{\#\pi_1(S_n)}\sqrt{\#\pi_2(S_n)}\sqrt{\#\pi_3(S_n)}. \tag{3.6}$$

Exercise 3.6. Verify each step above. Where was C-S inequality used? Why does $\chi^2_{\pi_j(S_n)}(x) = \chi_{\pi_j(S_n)}(x)$?

The product of three positive numbers certainly does not exceed the largest of these numbers raised to the power of three. It follows from this and 3.6 that

$$n = \#S_n \leq \max_{j=1,2,3} (\#\pi_1(S_n))^{\frac{3}{2}}.$$

We conclude by raising both sides to the power of $\frac{2}{3}$ that

$$\# \max_{j=1,2,3} \pi_j(S_n) \geq n^{\frac{2}{3}}$$

as claimed.

Exercise 3.7. Let Ω be a *convex* set in $\mathbb{R}^3$. This means that for any pair of points $x, y \in \Omega$, the line segment connecting x and y is entirely contained in Ω. Prove that $vol(\Omega) \leq \sqrt{area(\pi_1(\Omega))}\cdot\sqrt{area(\pi_2(\Omega))}\cdot\sqrt{area(\pi_3(\Omega))}$.

If you can't prove this exactly, can you at least prove using (3.6) and its proof that $\max_{j=1,2,3} area(\pi_j(\Omega)) \geq (vol(\Omega))^{\frac{2}{3}}$? This would say that a convex object of large volume has at least one large coordinate shadow. Using politically incorrect language this can be restated as saying that if a hippopotamus is overweight, there must be a way to place a mirror to make this obvious...

Exercise 3.8. (Project question) Generalize (3.6). What do I mean, you ask... Replace three dimensions by d dimensions. Replace projections onto two-dimensional coordinate planes by projections onto k-dimensional coordinate planes, with $1 \leq k \leq d-1$. Finally, replace the right hand side of (3.6) by what it should be...

Chapter 4

Graph theory and incidences

In this Chapter, we will give you a taste of some very useful ideas, which we will use heavily throughout the rest of the book. We will start off with some basic results from graph theory, and illustrate their use in incidence theory. Both areas will be the backbone of many of the results to follow.

1. Basic graph theory

Graph theory is a wide but powerful subject. In this section, we will give only the basics necessary to understand the content of the book. However, once you get comfortable with the ideas presented here, a little digging through the literature will lead you toward many rich and rewarding techniques.

A *graph* is a set of elements called *vertices*, and a set of pairs of vertices called *edges*. Vertices are normally represented by a set of points, and edges are represented by curves that connect pairs of points. In many cases, a given pair of vertices is either connected or not, so either there is a single edge connecting them or there is not. Graphs of this type are called *simple*. Sometimes, however, it is useful to consider *multigraphs*, or graphs where a pair of vertices may be connected by more than one edge. Some arguments in this book

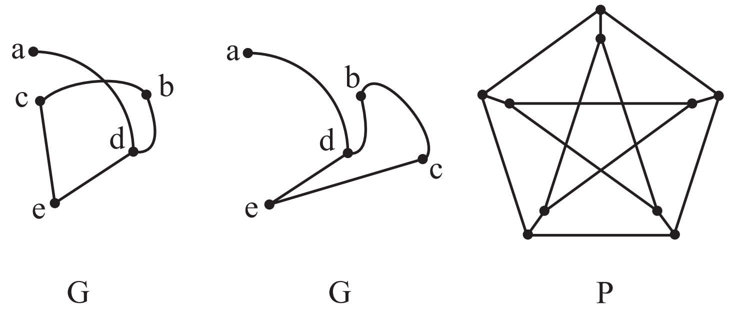


Figure 4.1. Two different drawings of the same graph G, and one drawing of the Peteresen Graph, P.

hinge on controlling the number of edges connecting a given pair of vertices, or its edge *multiplicity.*

If every vertex can be reached by every other vertex by traversing some number of edges, we will call the graph *connected.* This is not to be confused with *complete* graphs, where every vertex is directly connected to every other vertex directly. Also, for simplicity's sake, we will define edges to only occur between distinct vertices. This is merely a technicality, but it will simplify our calculations without any loss in generality for our needs.

The beginnings of graph theory were often concerned with *planar* graphs, or graphs whose edges need not cross. It is possible that a particular graph has been drawn in such a way that two edges cross, but it could be redrawn in such a way which retains all of the vertex connections, without any edges crossing. As usual, since we took the time to define planar, it seems as though there must be some graphs that are non-planar. If a graph is non-planar, then regardless of how we draw it, in order to preserve all of the connections between vertices, there must be some edges that cross each other. In order to get a feel for this, you should try to redraw P from Figure 4.1 without any crossings. What is the smallest number of crossings you can get? We define the *crossing number* of a graph, G, to be the minimum number

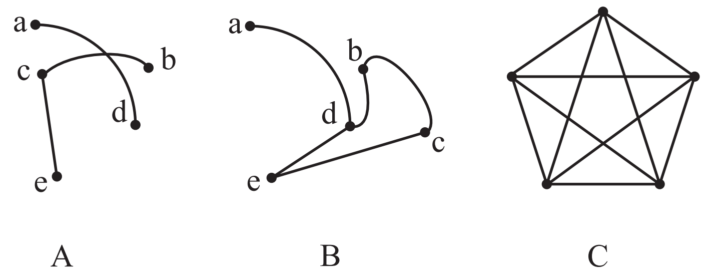


Figure 4.2. *A* is not connected. *B* is connected but not complete. *C* is both connected and complete.

of crossings that any redrawing of G has. We denote the crossing number of a graph, G, by $cr(G)$.

In order to get a hold of the crossing number of a graph, which is invariant under redrawings, we'll need to consider some of the other invariant properties of planar graphs. The first concept is that of a *face*. A face is any region bounded by edges. The graph G in Figure 4.1 has two faces. One face is the region contained by the edges connecting the following pairs of vertices: (b, c), (c, e), (e, d), and (d, b). The other face is the rest of the plane, or the outside of the previous face. By definition, there is always such a face outside of the graph. Now we can present some relationships that will always hold in a simple, planar graph. The following is called Euler's Formula.

Proposition 4.1. *Given a simple, connected, planar, graph, G, with n vertices, e edges, and f faces,*

$$n - e + f = 2.$$

Proof. One can easily derive this by induction on edges. Given any graph with one edge, we have two vertices and one face. If we wish to add another edge, we have to add another vertex, or we can connect to an existing vertex. If we connect to an existing vertex, we will generate another face. □

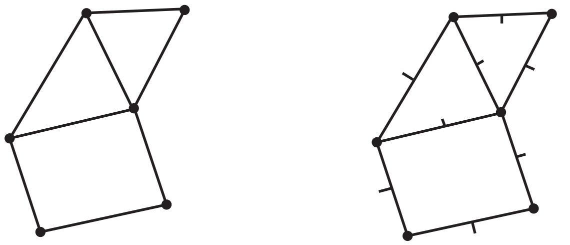


Figure 4.3. These are two pictures of the same graph, before and after drawing noses on the edges. It does not matter which side gets the nose, just that there is a way to tell one side from the other.

Proposition 4.2. *Given a simple, planar graph, G, with f faces and e edges,*

$$3f \leq 2e.$$

Proof. To see this, go through all of your edges and draw a nose on one side as shown in Figure 4.3. This will allow us to differentiate between two sides of an edge. If we count the number of total sides present in our graph, we will get $2e$, as each edge has two sides. Now count how many sides are present on each face. Each face requires at least three sides. So there are more than $3f$ sides of edges. □

Combining Propositions 4.1 and 4.2 yields the following useful corollary.

Corollary 4.3. *Given any simple, planar graph, G, with n vertices and $e > 2$ edges,*

$$e \leq 3n - 6. \tag{4.1}$$

Now we can get to the crux of our search, which is a simple reinterperetation of Corollary 4.3. Although the quantity "crossing number of G" doesn't immediately jump out of the inequality, it is hidden in the assumptions of the graph. In the above setting, G is planar, and therefore has no crossings. So if we have a *non*-planar graph, G_0, we know that the inequality will not hold. Suppose we look at a drawing of G_0 with the minimum number of crossings, and delete an edge that contributes at least one crossing. Now, we may not know where such an edge is in our graph, but we do know that if we delete any edge, that our number of edges will decrease by one. Call the resultant graph G_1, and then check to see if it is planar yet. How do we check if our graph is planar? See if it satisfies (4.1). Recall, this criterion depends only on the number of edges and vertices, so redrawing the graph will have no affect on the outcome. Using this method, we can remove edges until we satisfy (4.1). If we keep track of the number of edges that we have removed, we can have some idea how many crossings were present in the original graph, G_0. Notice that removing an edge can cause us to get rid of more than one crossing, so we will only have a lower bound on the number of crossings. This is made precise in the following theorem.

Theorem 4.4. *Given a graph, G, with e edges and n vertices, the crossing number is bounded below by:*

$$cr(G) \geq e - 3n + 6. \tag{4.2}$$

This relationship will give us a handle on the number of crossings in a graph without having to look too closely at the structure of the graph, which will prove to be quite useful.

2. Crossing numbers

The next Theorem is one of the most important tools in the book. We will use elementary probability theory alongside the basic graph theoretic results to prove it. You should make sure that no part of the proof is lost, as these ideas are very close to the center of this whole subject.

Theorem 4.5. *Let G be a graph with n vertices and e edges. If $e \geq 4n$, then*

$$cr(G) \gtrsim \frac{e^3}{n^2}.$$

To start, let us suppose we are given some graph G. By the Theorem 4.4 in the previous section,

$$cr(G) \geq e - 3n.$$

Choose a random subgraph, H, of G, by keeping each vertex with probability p, a number to be chosen later. What we mean here is that given all possible subgraphs of our graph, we can arrive at one subgraph in particular by keeping some of the vertices in the original graph, where each vertex is independently kept or thrown out.

Suppose that, independently, each of our vertices is chosen with some probability p. If we want to just consider the chosen vertices, we can consider a *subgraph*. It consists of vertices of the original graph corresponding only to the chosen vertices. If only one of the associated vertices of some edgehas been chosen, it will not be present in the subgraph, as an edge needs two vertices to make sense as we have defined them thus far. So, if one of these unfortunate edges that was considered in our original graph, but not in this particular subgraph, is removed, any crossings it contributed to the first graph will certainly not be present in our subgraph, no matter how it is redrawn.

As Figure 4.4 indicates, none of the edges associated with an unchosen vertex are in the random subgraph. So we will lose an edge if either vertex associated with that edge is not chosen. Further note that losing edges means that we may lose crossings.

Now, when we talk about the expected number of vertices, we mean that if we chose one of the possible subgraphs at random, we can expect some number of vertices. So the expected number of vertices in a random subgraph, with vertices chosen with probability p, will be np. Expected value is detailed in Appendix B.

So we have a handle on the expected number of vertices, but how many edges will remain? Well, we know that we have e edges to begin with, and each edge is kept in the subgraph only if both of its vertices are kept. What is the probability that an edge present in the original graph will be present in the subgraph? It will be the probability that both of its vertices are chosen. Since the vertices are chosen independently, each with probability p, the probability a given

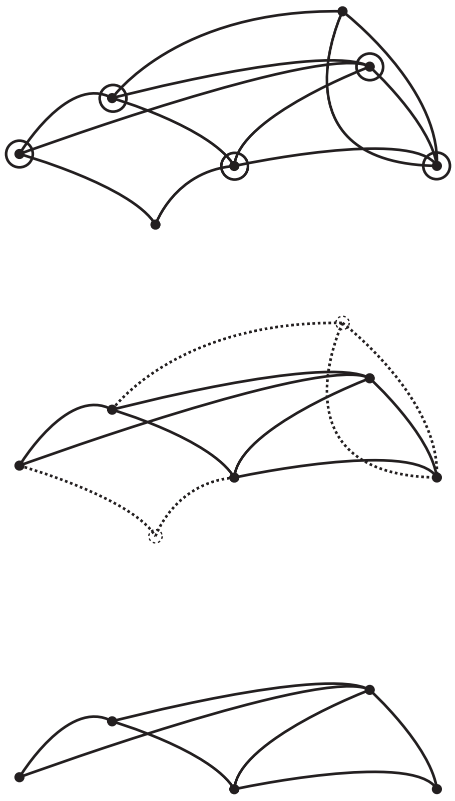


Figure 4.4. Suppose that the circled vertices from the top drawing of the graph were selected for a particular random subgraph. The middle shows the selected subgraph normally, with the doomed edges and vertices drawn as dotted lines. The bottom is a drawing of the selected subgraph.

edge will be chosen is p^2. So the expected number of edges in our subgraph will be ep^2.

We have only to figure out how many crossings we can expect, and then we can get to work cranking through the inequalities. Take note of how much reasoning must take place before you get to push symbols around. This just reinforces the point that the symbols are merely tools of for, and not the whole of mathematics. Since a crossing requires two edges that share no vertices, and each edge requires two distinct points, each crossing needs four points. So, crossings *may* occur with probability p^4, but recall that if an edge is not chosen, it could remove several crossings which were present in the original graph, but not in our subgraph. So we have an upper bound, not an exact value, for the expected number of crossings, $cr(G)p^4$.

So, in sum,

$$\begin{aligned} &\mathbb{E}(\text{vertices in } H) = np, \\ &\mathbb{E}(\text{edges in } H) = ep^2, \text{ and} \\ &\mathbb{E}(cr(H)) \leq cr(G)p^4, \qquad (4.3) \end{aligned}$$

where $\mathbb{E}$ denotes the expected value.

By (4.3) and linearity of expectation,

$$cr(G)p^4 \geq ep^2 - 3np.$$

Recall the strange condition in the statement of the Theorem, $e > 4n$. This is used to ensure that $\frac{4n}{e} < 1$, so it can be a probability. So, choosing $p = 4n/e$, as we may, since $e > 4n$, we obtain the conclusion of Theorem 4.5.

It might seem odd to make dedeuctive assertions using probabilistic ideas, but remember that we are not claiming that something is “highly likely” or that it will “probably happen”. We are making very careful statements that merely depend upon the calculated likelihoods of certain events. So do not worry, we are leaving nothing to chance!

We will now set graph theory aside for a little while, and introduce a closely related area, incidence theory. As you read through the next section, try to anticipate how we will apply graph theory to this setting, and then see how your ideas line up with the methods we describe. Remember, the more you put into this, the more stand to gain.

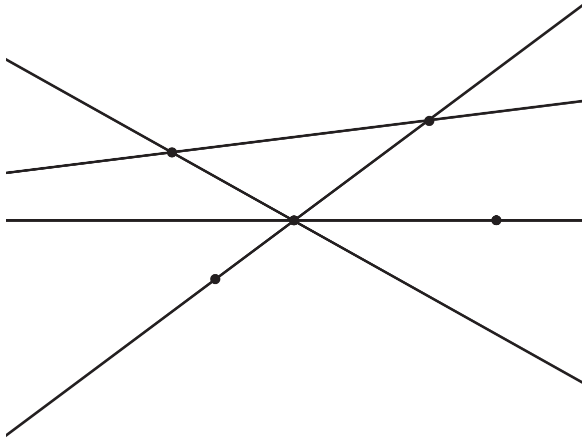


Figure 4.5. An example of four lines, five points, and nine incidences.

3. Incidence matrices and Cauchy-Schwarz

Let P be a finite set of n points in $\mathbb{R}^2$, and let L be a finite set of m lines. Define an *incidence* of P and L to be a pair $(p, l) \in P \times L : p \in l$. Let $I_{P,L}$ denote the total number of incidences between P and L. More precisely,

$$I_{P,L} = \#\{(p, l) \in P \times L : p \in l\}.$$

The following Figure has some points that lie on more than one line, as well as some lines incident to more than one point.

We already proved something about $I_{P,L}$ in Exercise 1.4, did we not? Let us think about it for a moment. Let $\delta_{lp} = 1$ if $p \in l$, and 0 otherwise. Then, by the Cauchy-Schwarz Inequality,

$$\begin{aligned}
I_{P,L} &= \sum_l \sum_p \delta_{lp} = \sum_l \sum_p (\delta_{lp} \cdot 1) \\
&\le \left(\sum_l \left| \sum_p \delta_{lp} \right|^2 \right)^{\frac{1}{2}} \left(\sum_l 1^2 \right)^{\frac{1}{2}} \\
&= \sqrt{m} \left(\sum_l \left(\sum_p \delta_{lp} \right) \left(\sum_{p'} \delta_{lp'} \right) \right)^{\frac{1}{2}} .
\end{aligned}$$

Notice that when we squared the second sum in p, we wrote it as the product of a sum in p and the same sum in p'. Our next step will be to seperate the case $p = p'$ from the case $p \neq p'$. This is a standard way of analyzing squared sums. Continuing,

$$\begin{aligned}
I_{P,L} &\le \sqrt{m} \left(\sum_l \sum_p \delta_{lp}^2 + \sum_l \sum_{p \neq p'} \delta_{lp} \delta_{lp'} \right)^{\frac{1}{2}} \\
&\le \sqrt{m} \left(mn + \sum_l \sum_{p \neq p'} \delta_{lp} \delta_{lp'} \right)^{\frac{1}{2}} .
\end{aligned}$$

Now, for each $(p, p') \in P \times P$, $p \neq p'$, there is at most one l such that $\delta_{lp}\delta_{lp'} \neq 0$. This is because $\delta_{lp} = 1$ means that $p \in l$, and $\delta_{lp'} = 1$ means that $p' \in l$. Since two points uniquely determine a line, the expression $\delta_{lp}\delta_{lp'}$ cannot equal to one for any other l. It follows that

$$\sum_l \sum_{p \neq p'} \delta_{lp} \delta_{lp'} \le \#\{(p, p') \in P \times P : p \neq p'\} = n(n-1).$$

Now it can be shown that the following theorem holds. You will explore the details in Exercise 4.2.

Theorem 4.6. *Let P be a set of n points in the plane, and let L be a set of m lines. Then $I_{P,L} \lesssim m\sqrt{n} + n\sqrt{m}$.*

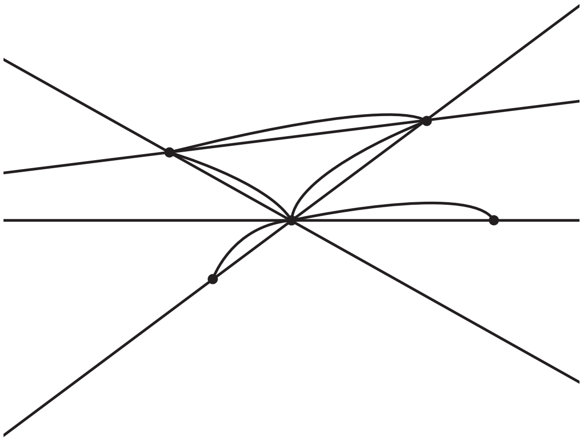


Figure 4.6. The same points and lines as before, but with their incidence graph drawn in as well.

4. The Szemeredi-Trotter incidence theorem

As pretty as this result is, it turns out that we can do better. The following improvement on Theorem 4.6 is due to Szemeredi and Trotter [**45**].

Theorem 4.7. *Let P be a set of n points in the plane, and let L be a set of m lines. Then $I_{P,L} \lesssim n + m + (nm)^{\frac{2}{3}}$.*

We now prove Theorem 4.7 using Theorem 4.5. In order to use Theorem 4.5, we construct the following graph. Let the points of P be the vertices of G, and let the line segments connecting points of P on the lines L be the edges. This exemplifies a technique that is extremely helpful in mathematics. We have a collection of objects that we want to know something about, so we model them in a setting where we can make some useful statements. Then we translate those statements back into our original setting, and if we constructed our model well, we might learn something new.

First, we need to show

$$e = I_{P,L} - m. \tag{4.4}$$

To see this, notice that each line will contribute as many edges incedences minus one. For example, a line with ten points on it needs only nine edges to connect all the points on that line by edges. So we lose one edge for the same reason every time we draw edges on a given line. So the number of edges must be the number of incidences minus the number of lines, as claimed in (4.4).

There are two possibilities. If $e < 4n$, then

$$I_{P,L} < m + 4n, \tag{4.5}$$

If $e \geq 4n$, then Theorem 4.5 kicks in, and we have

$$cr(G) \gtrsim \frac{e^3}{n^2} = \frac{(I_{P,L} - m)^3}{n^2}. \tag{4.6}$$

On the other hand, a crossing arises when two edges intersect at a point that is not in the set P, and therefore, not a vertex. Lines can only intersect each other once. There are m lines, which means that in total, lines can intersect each other at most $\binom{m}{2} \approx m^2$ times. Since edges are drawn along lines, edges certainly need not cross, except possibly when their related lines intersect. Therefore,

$$cr(G) \leq m^2.$$

If we compare the upper and lower bounds on the crossing number of our graph, we get

$$\frac{(I_{P,L} - m)^3}{n^2} \lesssim cr(G) \leq m^2.$$

This gives us another possible upper bound on $I_{P,L}$.

$$I_{P,L} \lesssim (nm)^{\frac{2}{3}} + m. \tag{4.7}$$

Combining (4.5) and (4.7), we obtain the conclusion of Theorem 4.7. The reason we can just add them is that even if one or the other dominates, surely their sum will dominate both.

At this point, make sure that you understand the construction of the graph G above. The specific kind of construction employed is a big part of this book. This Theorem's original proof was much more complicated. However, once it is viewed in a graph theoretic setting, it is quite simple.

One of the most misused words in mathematics is "sharp". Nevertheless, we are about to use it ourselves. We will show that Theorem 4.7 is sharp in the sense that for any positive integer n and m, we can construct a set P of n points, and a set L of m lines, such that

$$I_{P,L} \approx n + m + (nm)^{\frac{2}{3}}. \tag{4.8}$$

We shall construct an example in the case $n = m$, but we absolutely insist that you work out the general case in one of the exercises below. Let

$$P = \{(i,j) : 0 \leq i \leq k-1; 0 \leq j \leq 4k^2 - 1\}.$$

Let L be the set consisting of lines given by equations $y = ax + b$, $0 \leq a \leq 2k-1$, $0 \leq b \leq 2k^2 - 1$. Thus, we have n lines and n points. Moreover, for $x \in [0,k)$,

$$ax + b < ak + b < 4k^2,$$

and it follows that for each $i = 0, 1, \ldots, k$, each line of L contains a point of P with x-coordinate equal to i. It follows that

$$I_{P,L} \geq k \cdot \#L = \frac{1}{4}n^{\frac{4}{3}}.$$

Although Theorem 4.5 is quite powerful itself, if we explore what it says a little bit more, we can come up with a much stronger result, that will help us push beyond $n^{\frac{2}{3}}$.

Theorem 4.8. *Given a multigraph G with n vertices, e edges, and a maximum edge multiplicity of m, and $e > 5mv$,*

$$cr(G) \gtrsim \frac{e^3}{mn^2}.$$

This can be proven by repeatedly using probabilistic arguments similar to those used in the proof of Theorem 4.5. We will give you a sketch of the proof to follow in Exercise 4.9, but before this can make any sense, you must be absolutely clear and confident with the techniques we used there.

We will lean heavily on Theorem 4.8 for many results in this book. To quickly illustrate its power, here is a useful variant of the classical Szemerédi-Trotter Theorem, (Theorem 4.7).

Theorem 4.9. *Given n points and l curves in the plane, where no more than m of the curves go through any pair of points, and any two curves intersect one another at most c_0 times, for some finite constant, c_0, then the following upper bounds on $I(n,l)$, the number of point-curve incidences, and L_k, the number of curves with more than k points on them, hold.*

$$i) L_k \lesssim \frac{mn^2}{k^3} + \frac{mn}{k}$$

$$ii) I(n,l) \lesssim m^{\frac{1}{3}} (nl)^{\frac{2}{3}} + nm + l$$

You will prove this result in Exercise 4.10.

Exercises

Exercise 4.1. Why did Corollary 4.3 have fewer constraints on the types of graphs it could be used on than either of the two results preceeding it?

Exercise 4.2. Complete the details of the proof of Theorem 4.6.

Exercise 4.3. Restate, in your own words, why (4.3) is given as an inequality, and not an equality.

Exercise 4.4. For each n and m, construct a set P of n points and a set L of m lines, such that (4.8) holds. Use the argument in the case $n = m$ above as the basis of your construction.

Exercise 4.5. Let P be a set of n points in the plane. Let L be a set of m curves. Let $\alpha_{pp'}$ denote the number of curves in L that pass through p and p'. Let $\beta_{ll'}$ denote the number of points of P that are contained in both l and l'. Use the proof of Theorem 4.6 to show that

$$I_{P,L} \leq n\sqrt{m}\left(\sum_{p\neq p'} \alpha_{pp'}\right)^{\frac{1}{2}} + m\sqrt{n}\left(\sum_{l\neq l'} \beta_{ll'}\right)^{\frac{1}{2}}. \tag{4.9}$$

Exercise 4.6. Show that the estimate, $I(n) \leq Cn^{\frac{3}{2}}$, we just obtained for points and lines in the plane is best possible for points and lines in $\mathbb{F}_q^2$. *Hint*: Take all the points in $\mathbb{F}_q^2$ as your point set and take all the lines in $\mathbb{F}_q^2$ as your line set. If you are not familiar with vector spaces over finite fields, come back to this after reading Chapter 8.

Exercise 4.7. Show that the number of incidences between n points and n two-dimensional planes in $\mathbb{R}^3$ can be n^2. Suppose that we further insist that the intersection of any three planes in our collection contains at most one point. Prove that the number of incidences is $\leq Cn^{\frac{5}{3}}$.

More generally, prove that if we have n points and n $(d-1)$-dimensional planes in $\mathbb{R}^d$, then the number of incidences can be n^2. Show that the number of incidences is $\leq Cn^{2-\frac{1}{d}}$ if we further insist that the intersection of any d planes from our collection intersect at at most one point.

Exercise 4.8. Prove that n points and n spheres of the same radius in $\mathbb{R}^d$, $d \geq 4$, can have n^2 incidences. Use the techniques of this Chapter that when $d = 2$ the number of incidences is $\leq Cn^{\frac{3}{2}}$. What can you say about the case $d = 3$?

Exercise 4.9. Prove Theorem 4.8. First, delete edges independently with probability $1 - \frac{1}{k}$ and then delete all the remaining multiple edges–call this resulting graph G'. Calculate the probability p_e that a fixed edge e remains in G'. Now compare the expected number of edges and crossings in G' to the number in the original graph and use Theorem 4.5. Finally, use Jensen's inequality with $f(x) = x^a$, which says that $\mathbb{E}[x^a] \geq (\mathbb{E}[x])^a$ for $a \geq 1$.

Exercise 4.10. Prove Theorem 4.9. Use the the modified crossing number theorem, Theorem 4.8, and follow the proof idea of the classical Szemerédi-Trotter theorem, Theorem 4.7.

Exercise 4.11. Is Theorem 4.9 always stronger than the one given by Exercise 4.5? Give explicit examples to support your belief.

Chapter 5

The $n^{4/5}$ theory

In this chapter, we shall use graph theory that already bore fruit in the previous chapter to improve the Erdős exponent from 2/3 to 4/5. The key new feature here is the use of bisectors. We shall take advantage of the fact that the centers of circles passing through a given pair of points lie on the bisector line.

1. The Euclidean case: straight line bisectors

Suppose that a set, P, of n points determined t distinct distances. Draw a circle centered at each point of P containing at least one other point of P. By assumption, we have at most t circles around each point, and thus the total number of circles is nt. By construction, these circles have $n(n-1)$ incidences with the points of P. The idea now is to estimate the number of incidences from above in terms of n and t and then derive the lower bound for t.

Delete all circles with at most two points on them. This eliminates at most $2nt$ incidences, and since we may safely assume that t is much smaller than n, the number of incidences of the remaining circles and the points of P is still $\gtrsim n^2$. Form a graph whose vertices are points of P and edges are circular arcs between the points. This graph G has $\approx n$ vertices, $\approx n^2$ edges, and the number of crossings is $\lesssim (nt)^2$.

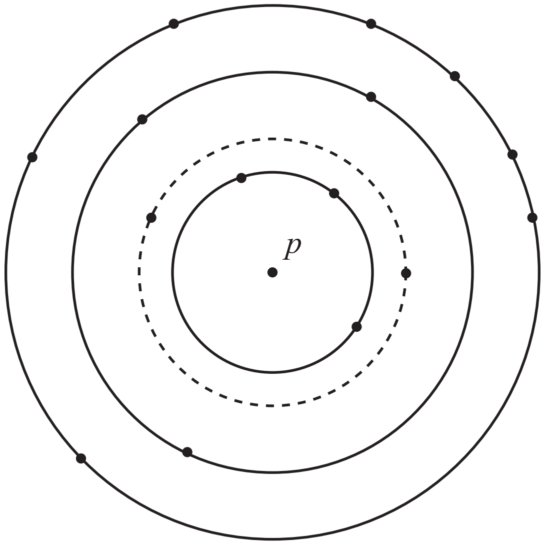


Figure 5.1. Edges along arcs of circles contributed by a point, p, with one of its circles deleted.

Suppose for a moment that we can use Theorem 4.5. Then

$$\frac{e^3}{n^2} \lesssim cr(G) \lesssim (nt)^2,$$

and since $e \approx n^2$, it would follow that

$$n^4 \lesssim n^2 t^2,$$

which would imply the Erdős Distance Conjecture. Unfortunately, life is harder than that since Theorem 4.5 only applies if there is at most one edge connecting a pair of vertices. In our case, we may assume that there are at most $2t$ edges connecting a pair of vertices (why? see Exercise 5.1 below). Applying Theorem 4.8 we see that

$$\frac{e^3}{tn^2} \lesssim cr(G) \lesssim n^2 t^2,$$

which implies that

$$t \gtrsim n^{\frac{2}{3}},$$

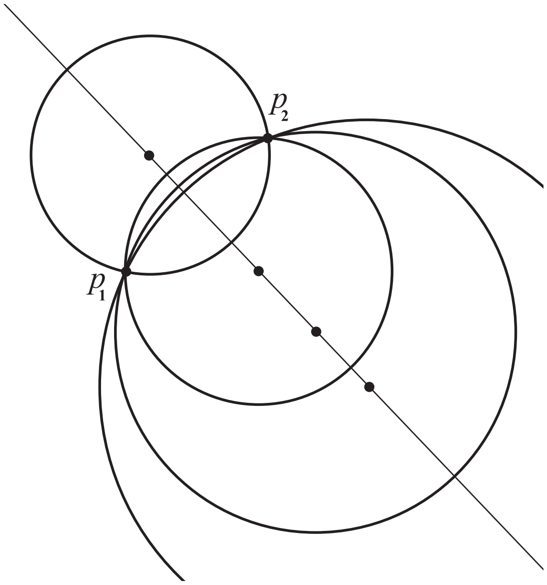


Figure 5.2. The bisector of p_1 and p_2 has four points on it. The arcs of the circles centered at those four points could contribute as many as four edges between p_1 and p_2.

Moser's bound from Chapter 2. All of this for $n^{\frac{2}{3}}$?! We must be able to do better than that! How can we possibly hope to improve the estiamate? One way is to study edges of high multiplicity separately.

We try to take advantage of the following phenomenon. Let $p, p' \in P$. The centers of all of the circles that pass through p and p' are located on the bisector, $l_{pp'}$, of the points p and p' in P [1].

Let us consider all of the bisectors with at least k points on them. How many such bisectors are there? Recall that the Szemeredi-Trotter incidence bound (Theorem 4.7) says that the number of incidences between n points and m lines is $\lesssim (n + m + (nm)^{\frac{2}{3}})$. Let m_k denote the number of lines with at least k points. Then the number

[1] The *bisector* of p and p' is the set of points that are equidistant to p and p'. Formally, $l_{pp'} = \{z \in \mathbb{R}^2 : |z - p| = |z - p'|\}$. In the Euclidean metric, this turns out to be the line perpendicular to $\overline{pp'}$ through their midpoint. For more general metrics see Exercise 5.2.

of incidences is at least km_k. It follows that

$$km_k \lesssim n + m_k + (nm_k)^{\frac{2}{3}},$$

and we conclude that

$$m_k \lesssim \frac{n}{k} + \frac{n^2}{k^3}. \tag{5.1}$$

Consider the following Lemma.

Lemma 5.1. *The number of incidences of lines with at least k points is* $\lesssim \frac{n^2}{k^2} + ctn\log n$.

This implies that bisectors with at least k points on them have

$$\lesssim n\text{log}(n) + \frac{n^2}{k^2} \tag{5.2}$$

incidences with the points of P.

Let P_k denote the set of pairs, (p, p'), of P connected by at least k edges. Let E_k denote the set of edges connecting those pairs. Each edge in E_k connecting a pair, (p, p'), corresponds to exactly one incidence of $l_{pp'}$ with a point, p'', in P. However, an incidence of such a p'' with some $l_{pp'}$ corresponds to at most $2t$ edges in E_k, since there at at most t circles centered at p''. It follows that

$$\#E_k \lesssim tn\log n + \frac{tn^2}{k^2}.$$

Note that we are almost certainly over counting E_k here, since we are removing all possible edges corresponding to incidences–not just those that contribute to high multiplicity.

Now, if we choose $k = c\sqrt{t}$, for an appropriate constant c, then

$$\#E_k \leq \frac{n^2}{2}.$$

If we now erase all the edges of E_k, there are still more than $\frac{n^2}{2}$ edges remaining. Applying Theorem 4.8 once again, we see that

$$\frac{e^3}{kn^2} \leq cr(G) \leq n^2t^2.$$

Since $k \approx \sqrt{t}$ and $e \approx n^2$, it follows that

$$t \gtrsim n^{\frac{4}{5}}.$$

Pending the proof of Lemma 5.1, we have just proved the following theorem of Szekely [**44**].

Theorem 5.2 (Székely [**44**]). *Let P be a set of n points in the plane. Then*

$$\#\Delta(P) \gtrsim n^{\frac{4}{5}}.$$

Now to prove Lemma 5.1

Proof. First, notice that the number of lines incident to 2^i points is at most $c\frac{n^2}{2^{3i}}$, provided $2^i \leq \sqrt{n}$. This is because if there were more, the total number of such lines would exceed the bound from Theorem 4.7, part (a). You will work out the details for this in Exercise 5.4.

This takes care of the lines incident to fewer than $\sqrt{n}$ points. Since a line with fewer than $\sqrt{n}$ points can contribute no more than k incidences, we get fewer than $\frac{n^2}{k^2}$ incidences from these lines. If a given line is incident to more than $\sqrt{n}$ points, the Szemerédi-Trotter theorem will no longer help. This case is even easier though, in light of a simple *inclusion-exclusion*[2] argument in [**43**]. Since lines can intersect each other at most only once, by definition, we're guaranteed that there can only be so many lines incident to a relatively large number of points. After recognizing this, there are merely a few simple things to count and we are done.

To nail down the inclusion-exclusion argument, call each line incident to more than $l \geq \sqrt{2n}$ points A_i, and let $|A_i|$ be the number of points incident to that line. Recall N_l is the number of lines with between l and $2l$ points, where $l \geq 4\sqrt{n}$. For the lemma to hold, we need $N_l \leq \frac{4n}{l}$. So given $l \geq \sqrt{n}$, let us suppose that $N_l \geq \frac{2n}{l}$, and arrive at a contradiction.

$$n = |E| \geq \left| \bigcup_{l \leq |A_i| \leq 2l} |A_i| \right| \geq \sum_{i=1}^{N_l} \left| A_i \setminus \left(\bigcup_{j=1}^{i-1} A_j \right) \right|,$$

[2] Inclusion-exclusion refers to statements like the following: $|A \cup B| = |A| + |B| - |A \cap B|$.

upon possibly reordering the A_i's to put those considered in the union first. This sum is clearly greater than or equal to

$$\sum_{i=1}^{N_l} \max(0, m-i) \geq \sum_{i=1}^{\frac{4n}{l}} \max(0, m-i) \geq \sum_{i=1}^{\sqrt{n}} \max(0, 4\sqrt{n}-i) \geq$$
$$\geq \sqrt{n}(4\sqrt{n}-\sqrt{n}) \geq 3n.$$

So we have a contradiction, implying that $N_l \leq \frac{4n}{l}$.
Now, to get the total number of incidences, A_i, we'll sum over all of them. However, when doing so, we will group lines by which powers of two are directly greater than and less than the number of points on each line.

$$\sum_{i:|A_i|\geq|\sqrt{n}} 2|A_i| \leq \sum_{i:2^j \leq i \leq 2^{j+1}} 2^{j+1} 2N_{2^j} \leq \sum_{i:2^j \leq i \leq 2^{j+1}} 2^{j+1} 2\frac{4n}{2^j} \leq$$
$$\leq 4n \sum_{\sqrt{n} \leq 2^i \leq n} 1 \leq 4n \log n.$$

This completes the proof of the Lemma 5.1. □

Another important idea is illustrated in the previous proof. When seeking to bound something like this, it is useful to consider different cases. Above, we had different bounds for lines with "many" and "few" points. (Exercises 5.3 and 5.4 illustrate how the lines with many and few points are bounded differently.) We found a balance, and we gained over either estimate by using both. This is of course hidden in the fact that the upper bound in Theorem 4.7 has all of the possible dominating terms summed together, so it handles all cases simultaneously. One could just as easily state the theorem as:

Theorem 5.3. *Let P be a set of n points in the plane, and let L be a set of m lines. Then at least one of the following is true:*

$$i) I_{P,L} \lesssim n,$$
$$ii) I_{P,L} \lesssim m, \textit{ or}$$
$$iii) I_{P,L} \lesssim (nm)^{\frac{2}{3}}.$$

We heavily exploit the fact that we can address the different bounds seperately, and that is how we gain over the $n^{\frac{2}{3}}$ bound we achieve at first.

2. Convexity and potatoes

Throughout the book so far, we have asked you to pause after some of the main arguments, and think about what aspects of the standard Euclidean metric were really necessary to apply the techniques that made things work. In this section we take only a slight diversion from that general scheme by introducing a new class of metrics that does not contain the standard Euclidean metric. In this case, it will be necessary to work a little harder, and build on the ideas already presented, rather than deduce possible relaxations to previous assumptions.

The metrics we introduce here are called *potato* metrics. The classification of such metrics is that they are *strictly convex*, which we will describe below, and all pairs of their bisectors can intersect in at most c_0 points, for some constant c_0. As mentioned in Chapter 1, these need not be symmetric.

What does it mean that a metric is strictly convex? One way of visualizing a strictly convex metric is to pick a point, and draw a "circle" around it, corresponding to all the points of some fixed distance from that point. If these points form a strictly convex shape, then we will call our metric strictly convex. Basically, strictly convex excludes flat sides, where merely convex would allow for such flat sides. A more precise definition of convexity is that for any two points in a set, any *convex combination* of those two points is in the set as well. A convex combination of two elements, a and b is $\lambda_1 a + \lambda_2 b$, where λ_1, and $\lambda_2 \in \mathbb{R}$, and $\lambda_1 + \lambda_2 = 1$. In a strictly convex set, any convex combination of points can not lie on the *boundary* or outermost points of the set.

What kinds of metrics could be convex but not circular? What if you were in a canoe, floating down a river. You could measure the "distance" between two points as the time it takes to get from one to another in your canoe. It will probably take less time to flow with the current of the river than against it, so all of the points that you can reach in ten seconds that are more or less downstream from your boat are probably farther away from your current position than the points

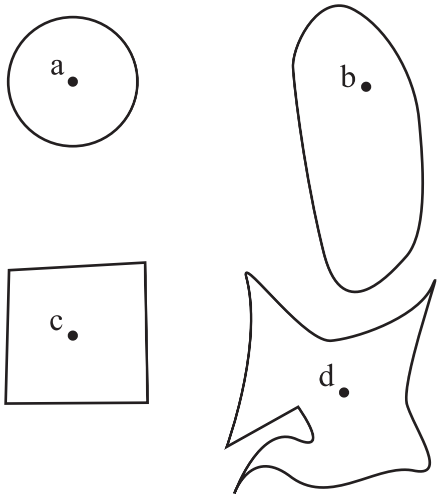


Figure 5.3. The circle centered at a is from the standard Euclidean metric. The circle centered at b is a strictly convex metric, could be thought of as from a canoe metric. The circle centered at c is from a convex, but not stricly convex metric, and the locus of points centered at d cannot be the circle of any metric, as they do not form a convex shape.

that you could reach in ten seconds that are relatively upstream from where you are right now.

If you look at Figure 5.3, you should have a pretty good idea of what is strictly convex and what is not. Now to address the issue of bisectors. In the Euclidean case, bisectors of points were just straight lines that ran through the midpoint of two points, and were perpendicular to the line through the two points. However, if we have two (or more) pairs of points, where the lines through each point pair are parallel, and the midpoints of each point pair all lie on the same line perpendicular to the parallel lines, we will have the same bisector for each point pair! This means that there are clearly more than constantly many intersections between different bisectors. So the standard Euclidean metric is not a potato metric.

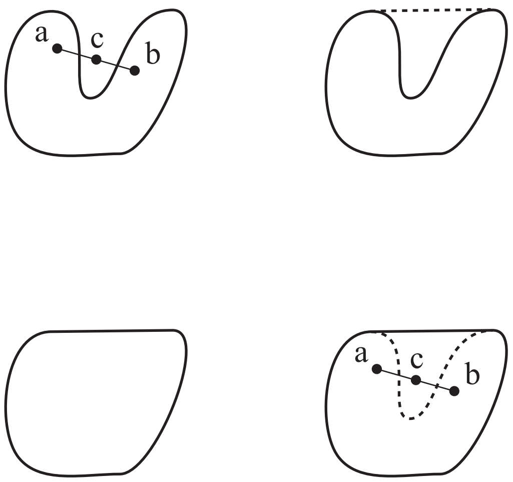


Figure 5.4. The first picture is a non-convex set, V, as illustrated by the fact that $c = \frac{1}{2}a + \frac{1}{2}b$, a convex combination of two points, $a, b \in V$, is **not** in V. The second picture illustrates how we find the convex hull, and the third picture is the convex hull of V. The last picture shows that the point c, from before, is indeed contained in V.

Since we are dealing with convexity so directly here, it is a good time to introduce the notion of *convex hull*. Suppose you have a set $V \in \mathbb{R}^d$, that is not necessarily convex. The convex hull of V would be the set of all convex combinations of elements in V. You can think of this as "filling in the gaps" of a non-convex set to construct a convex set. This is an extremely important notion in mathematics. We do not use it much in this particular book, but if you continue to study mathematics, you will find that it pops up all over the place!

Notice that convex sets obviously contain their interiors. A circle is just a closed curve in the plane, but a disk contains all of the points inside of that closed curve. Be careful of the distinction between convex sets, and the convex curves that form their boundaries. In order to make this more precise, we will introduce the symbols ∂

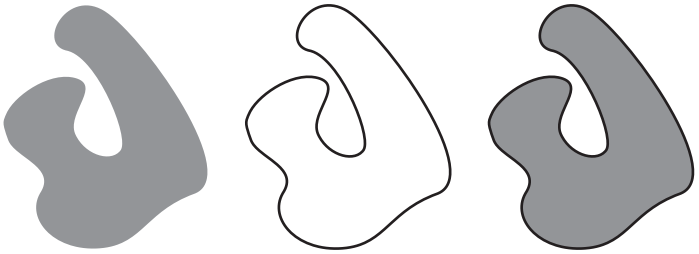


Figure 5.5. The first picture is a set J, which could also be called $(\partial J)^\circ$, the interior of the boundary of J. The second picture is the boundary of J, denoted ∂J. The third is $J \cup \partial J$.

and °. This is illustrated in Figure 5.5. If we have a set S, let the outer-most points be called the boundary, and denoted ∂S. All of the points contained properly inside of the boundary, and not containing the boundary will be denoted S°. Much, much more can be said about topology and the theory of open and closed sets, but we make no attempt to address that here. Suffice it to say that we will just borrow some notation.

Now that we have the general concepts of convex hull, boundary, and interior, we can present the following lemma, which will also serve to illustrate another way in which strictly convex metrics exhibit their ability to behave well. It may seem a bit unusual at first, but it will become quite handy when we try to deal with any strictly convex metrics, and in the section to come, potato metrics. If we are considering the metric K, we will call all distances with respect to this metric *K-distances*, and similarly, refer to *K-circles*.

Lemma 5.4. *Given a strictly convex metric, K, two K-circles, $C_K(x, r)$ and $C_K(y, s)$ can intersect in at most two points, assuming that $x \neq y$, and $r \neq s$, where x and y are points in the plane, and r and s are the radii of the corresponding K-cricles centered at x and y.*

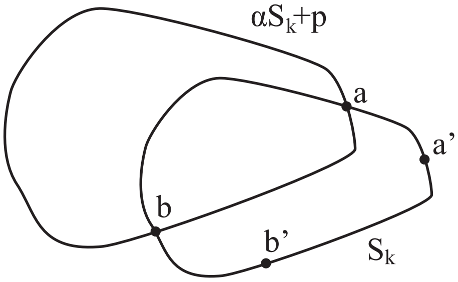


Figure 5.6. Here is a picture of S_k intersecting with the translated and dilated $\alpha S_k + p$. Note that the points a and b lie on both circles, and in some sense, a' and b' lie on S_k where a and b lie on $\alpha S_k + p$.

Proof. Without loss of generality, we will consider the second K circle to have radius 1, and be centered at the origin, or that $y = (0, 0)$, and $s = 1$. This is perfectly acceptable, because once we have a result in that case, we are free to *translate*[3] and dilate any other situation into this one. We will call the radius 1 K-circle centered at the origin S_K. The other one will be called $\alpha S_K + p$ for appropriate $\alpha > 0$ and p. You will do an example of finding such α and p in Exercise 5.8.

We will continue this proof by way of contradiction. Suppose that a, b, and c are three distinct points that lie on both K-circles. We can assume that they are not collinear, as this would immediately violate our strict convexity assumption. We know that $a, b, c \in S_K$, this also means that $a', b', c' \in S_K$, where $a' = \alpha^{-1}(a - p)$, $b' = \alpha^{-1}(b - p)$, and $c' = \alpha^{-1}(c - p)$. Call D_K the set of points on and inside of S_K, (so S_K is the boundary of D_K, ∂D_K, and D_K is the interior of S_K, S_K°). Let T and T' denote the triangles abc and $a'b'c'$, respectively. Call D'_K the convex hull of $T \cup T'$. Since a, b, c, a', b', and c' are all in D_K, and D_K is convex, $D'_K \subset D_K$. Since they all lie on S_K, they must also all lie on $S'_K = \partial D'_K$.

[3] If you are unfamiliar with it, the notion of translation is explored in and around Proposition 10.1. The context there is in vector spaces over finite fields, but the proof reads nearly identically for $\mathbb{R}^d$ with a few obvious modifications.

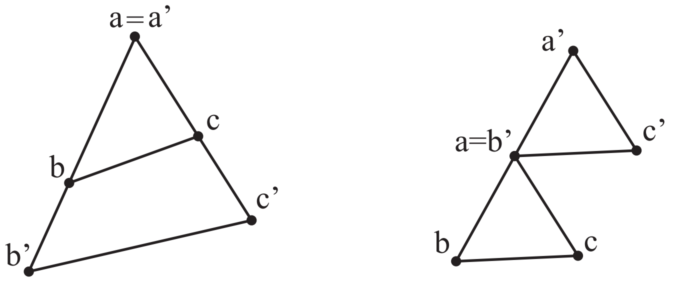


Figure 5.7. These are some of the possibile ways $a = a'$ and $a = b'$ could occur.

Observe that S'_K consists of some number of edges of the triangles T, T', and at most two additional line segments. You will work these details out in Exercise 5.7. We will handle two seperate cases. Suppose for now, that for each triangle, at most one of the pair of congruent edges is in S'_K, (e.g. either ab or $a'b'$ is in S'_K, but not both). The boundary of S'_K consists of as many as, but no more than, five line segments. So it can have no more than five vertices.[4] If all six of the aforementioned points were on S'_K, then at least three of them must be collinear.

Now, if the three collinear points are distinct, then, since they were all on S_K, that means that S_K contains a line segment. This violates the strict convexity condition of our metric, so the three collinear points cannot be distinct. That means that, again, without loss of generality, either $a = a'$, or $a = b'$. We say "without loss of generality" here because if it were actually the case that $c = c'$, we could simply rename the points so that c was a, and continue the proof precisely as written. The same would hold if $b = b'$.

If $a = a'$, then $\alpha \neq 1$, which means that a, b, and, b' are distinct and collinear. If $a = b'$, then a, a', and b are distinct and collinear. Either way, we argue as in the last paragraph to get a contradiciton.

[4] Here, by vertices we refer to the corners of a shape, not the vertices of a graph.

Now, recall our earlier assumption, that for each triangle, at most one of the pair of congruent edges is in S'_K. If that was not actually the case, then we have a scenario where, without loss of generality, the segments ab and $a'b'$ are contained in S'_K. Again, this means that there are at least three distinct collinear points in S'_K, which violates our strict convexity assumption.

So, we have exhausted all possible cases of three points on the intersection of three or more points in $C_K(x, r)$ and $C_K(y, s)$, and rigorously shown the result. □

3. Székely's method for potato metrics

This section is quite dense. We include it here because it is basically the same argument as above, but with some minor modifications. These modifications get quite involved, and if you start to lose sight of the goal, feel free to start the next chapter and come back to this section later.

We presented Székely's method above, which gave us $n^{\frac{4}{5}}$ for the Euclidean metric. We will now set out to get a lower bound on the size of distance sets of potato metrics. In the proof to follow, as in many such proofs, when we wish to show something about all objects in a particular class, we will pick an arbirary member of that class, and show that the desired result holds. Then we know it is true for any element in that class. With this in mind, we fix an arbitrary potato metric, K, and proceed.

The basic idea behind this argument is the same as behind the proof of Theorem 5.2. We draw K-circles about each point, such that they cover all of the points of the set. Define t as before. Again, we will delete all circles with strictly fewer than three points on them. We can get away with this for the same reasons that we got away from it last time. We will construst the same kind of multigraph, G, using the points as vertices, and the arcs of the K-circles connecting consecutive points as edges. Since we have no more than t K-circles around any given point, there are about n^2 edges in our graph.

So far, everything is the same, but when we try to get upper and lower bounds for the crossing number with Theorem 4.8, we have a problem. Since the bisectors of potato metrics are not necessarily straight lines, we do not, a priori, have a good upper bound on maximum edge multiplicity.

We know that if there are k edges connecting two vertices, that there must be k points on the bisector of the pair of points corresponding to the pair of vertices connected by so many edges. So we will eventually use Theorem 4.9 to get good bounds on edge multiplicity, as before. In order to use that result, though, we will need to know how many bisectors can go through a pair of points.

So let us start by constructing a new multigraph, H, with the same vertices as G, but whose vertices are arcs of the $\binom{n}{2}$ different bisectors. However, in this graph, we will actually make a small adjustment to the edges. If a bisector is incident to a point that does not contribute any edges to G, we will edit the corresponding edge in H by drawing it in such a way as to circumvent the point, but not disturb the edge crossings in any way. This is illustrated in Figure 5.8. We can do this as points are infintesimally small, so we can make corrections to the edges that are smaller than any distance between any point and any K-bisector.

Also, we know that the bisectors are distinct because any two bisectors can intersect only finitely many times, by definition of the potato metric. Of course we will not consider the arcs that go out past all of the points to infinity. Keep in mind that we are doing this to get a handle on the maximum edge multiplicity, m, of H.

Proposition 5.5. *If K is a potato metric, then the maximum edge multiplicity, m of H, the graph of K-bisectors determined by a set of n points is at most $2t$, where t is the maximum number of K-circles around any point.*

Assuming the Proposition for now, we can appeal to Theorem 4.9 and get that the number of bisectors in H that contain at least k points is bounded above by $\frac{tn^2}{k^3}$ as long as $k \lesssim \sqrt{n}$. Similarly, the number of bisectors containing at least k points is bounded above by $\frac{tn}{k}$ when $k \gtrsim \sqrt{n}$. So, as before, if we remove all edges of multiplicity greater than k, the most edges we will lose will be bounded above by:

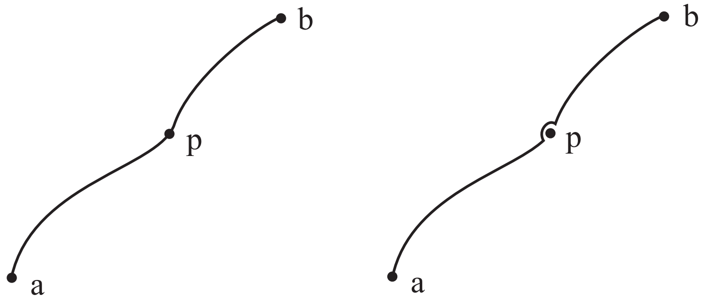


Figure 5.8. If the point p does not contribute edges to G that pass through the K-bisector shown, it is unnecesary to consider it in H. The figure on the left shows two edges, connecting a to p, and p to b. The figure on the right shows only one edge, connecting a and b.

$$\sum_{\imath : k < 2^i \lesssim \sqrt{n}} \underbrace{\frac{tn^2}{2^{3i}}}_{\text{bisectors}} \underbrace{2^i}_{\text{arcs}} + \sum_{\imath : \sqrt{n} \lesssim 2^i \lesssim n} \underbrace{\frac{tn}{2^i}}_{\text{bisectors}} \underbrace{2^i}_{\text{arcs}} \lesssim \frac{tn^2}{k^2} + tn\log_2 n.$$

Again, we can let $k \approx \sqrt{t}$, and still retain about n^2 edges after deleting edges with multiplicity greater than k. Now when we apply Theorem 4.8, we have the same upper and lower bounds as before:

$$\frac{n^6}{t^{\frac{1}{2}} n^2} \lesssim \frac{e^3}{kn^2} \lesssim cr(G) \lesssim n^2 t^2.$$

After doing the arithmetic, we get

$$t \gtrsim n^{\frac{4}{5}}.$$

This means that we have shown the following theorem, pending proof of Proposition 5.5.

Theorem 5.6. *Let P be a set of n points in the plane. Suppose the metric used to measure distance is a potato metric, that is, that it is* *strictly convex, and that all pairs of bisectors can intersect each other in at most c_0 points, where c_0 is some constant. Then*

$$\#\Delta(P) \gtrsim n^{\frac{4}{5}}.$$

In order to prove Proposition 5.5 about H, that $m \lesssim t$, we will need to look at the way that K-circles intersect. Now the lemma in the previous section does not look as strange! Indeed, we need Lemma 5.4 to start us off. We will use this to prove the following lemma, which will be the final step before we can set off proving Proposition 5.5. The following proofs will most likely require several readings for all of the ideas to become apparent. These are highly technical arguments, so do not worry if something seems unclear the first time through.

Lemma 5.7. *If $C_K(x,r)$ and $C_K(y,s)$ intersect in two points. Let x_o and x_e be the points on $C_K(y,s)$ with, respectively, the largest and smallest K-distances to y. Then each of the intersections will lie on different sides of the line l, which passes through both x_o and x_e.*

Proof. We can assume $x \neq y$, as in this case, there are no intersections unless $r = s$, in which case there are infinitely many intersections. So either way, that case violates our assumption of only two intersections.

We will first notice that by definition of x_o, it is unique. We know it is unique because it is the intersection of $C_K(x,r)$ and l, which is a single point. Let A_+ denote the arc above x_o and x_e, and A_- denote the arc below. Since x_o is unique, there is at least one intersection, on each of A_+ and A_-, close to x_o. So for every s'' strictly between $\|y - x_o\|$ and $\|y - x_e\|$, there is exactly one intersection with A_+ and A_-, because if there were more than two intersections, it would violate the strict convexity assumption by Lemma 5.4. Of course, there is once again only one unique intersection between $C_K(x,r)$ and $C_K(y, \|y - x_e\|)$, which, by definition occurs at x_e.

□

Finally, we can prove Proposition 5.5. You should go back and reread the construction of the graph H for this proof to make sense. Consider the K-circles $C_K(x,r_i)$ and $C_K(y,s_j)$. By the criterion given

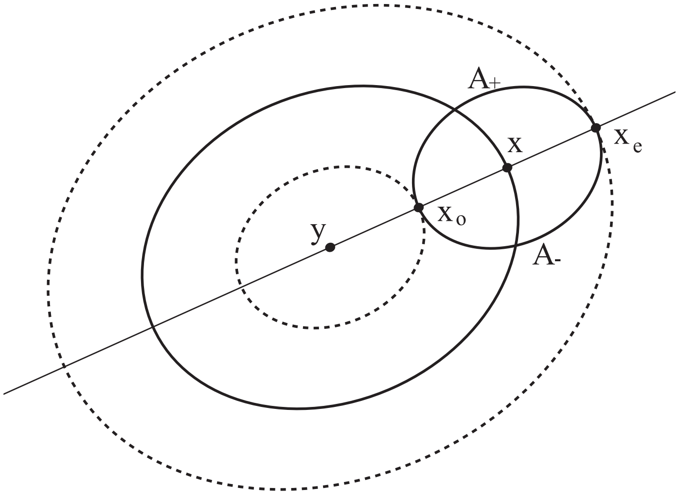


Figure 5.9. This is one possible depiction of $C_K(x,r)$ and $C_K(y,s)$ intersecting in exactly two points.

for an edge to hit a point corresponding to the K-bisector it lies on, we need arcs from a_{ij}[5] to b_{ij} that are edges on both $C_K(x,r_i)$ and $C_K(y,s_j)$. We aim to show that this can happen at most $2t$ times, by showing that only two of the t possible pairs satisfy the requisite conditions to have an edge in H hit x.

Now, if the arc between a_{ij} and b_{ij} corresponds to an edge in G, it contains either x_o or x_e. Without loss of generality, we will suppose that it contains x_o. This means that any other arc between $a_{ij'}$ and $b_{ij'}$, for $j' < j$ cannot be in G, as it would split the edge connecting a_{ij} and b_{ij}. There is a similar argument for x_o and $j' > j$.

So, each circle about x can contribute at most 2 edges in H, and there are no more than t circles about x. Therefore, our maximum edge multiplicity in H is $2t$, as claimed.

[5] Here, the subscript ij denotes two seperate indices, not a product of $i \times j$.

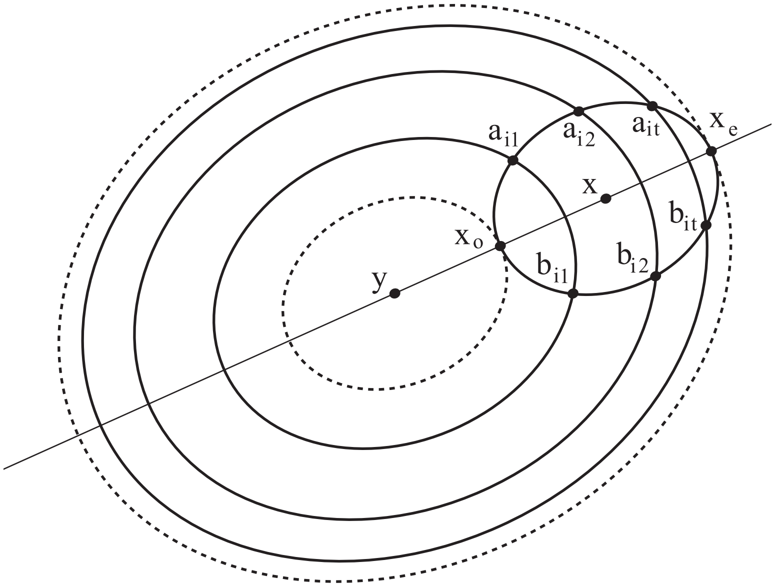


Figure 5.10. Here, we show only a fixed radius, r_i, for the K-circle centered at x.

Exercises

Exercise 5.1. Explain why there can be at most $2t$ edges connecting two vertices in the graph G from the proof of the Euclidean setting. Think about where the edges come from, and derive a contradiction if there are more than $2t$ edges connecting two vertices.

Exercise 5.2. Consider the l_1 metric defined in Exercise 0.4. Try to figure out what bisectors look like for this metric. Look at the following point pairs first: (1,0) and (-1,0), then try (0,0) and (1,2), and finally examine (1,1) and (-1,-1). Why was the last example so different?

Exercise 5.3. Explicitly show why there can be no more than $c\frac{n^2}{k^3}$ lines with more than k points on them, when $k \lesssim \sqrt{n}$.

Exercise 5.4. After doing Exercise 5.3, what can you say about lines with more than k points on them, when $k \gtrsim \sqrt{n}$. It is important to understand how these bounds are different, and what the plus signs in the right hand side of Theorem 4.7 mean.

Exercise 5.5. Show that any convex set is its own convex hull.

Exercise 5.6. Suppose V is any set with a concavity, that is, a convex combination, c, of two points in V that is not itself in V. Show that the convex hull of V is not strictly convex. *Hint*: You might want to distinguish points on the boundary of sets from points not on the boundary of sets.

Exercise 5.7. Convince yourself that if $a, b, c, x \in \mathbb{R}^2$, $\alpha > 0 \in \mathbb{R}$, $a' = \alpha(a + x)$, $b' = \alpha(b + x)$, and $c' = \alpha(c + x)$, then for triangles $T = abc$ and $T' = a'b'c'$, the convex hull of $T \cup T'$ is a polygon with at most five edges, or a line segment. *Hint:* Notice that at most one of the segments ab or $a'b'$ can be on the boundary of the convex hull.

Exercise 5.8. Use the statement that we showed precisely in Lemma 5.4, (the "WLOG" statement, for $y = (0, 0)$, and $s = 1$), to show that for a strictly convex metric K, the following K-circles can intersect at most twice: $C_K((0, 2), 2)$ and $C_K((2, 0), 2)$.
Note that we do not specify the metric, K, as we do not need to. You will have to pick one of the circles to translate to the origin, and then translate both accordingly. Do this with a change of variables. If you let $y = (2, 0) - (2, 0) = (0, 0)$, then you will have to let $x = (0, 2) - (2, 0) = (-2, 2)$, where the subtraction here denotes vector, or coordinatewise, subtraction. Then you will have to do something similar to get the associated K-radius of the K-circle centered at y to be 1. Just as a heads up, it is not a simple as just dividing both radii by two. Why?

Exercise 5.9. In the proof of the potato metric theorem, Theorem 5.6, we needed an estimate for the maximum edge multiplicity of the graph, H, consisting of points in the plane and potato metric bisectors. Lemma 5.5 provides a sharp bound, but without going

through all of that, prove that $m \lesssim t^2$ using the following two facts. The term *K-radius* refers to the K-distance a point on a K-circle is from the center of the K-circle.
1) Given two points, x and y, we know how many K-circles can maximally be centered at each on of them.
2) It takes two points on circles of the same K-radius to determine a bisector.

Exercise 5.10. Show that the intersection of two convex sets is convex. *Hint*: All you need to do is write down the definition of convexity, and the definition of intersection.

Chapter 6

The $n^{6/7}$ theory

In this chapter, we present the beautiful Solymosi-Toth argument that will get us up to $n^{6/7}$, which opens the door to further important developments that we sketch in the next chapter. We start out with the following beautiful observation due to Jozsef Beck [**3**]. The proof we give is from [**40**].

1. The setup

Lemma 6.1. *Let P be a collection of n points in the plane. Then one of the following holds:*

(1) *There exists a line containing $\approx n$ points of P.*

(2) *There exist $\approx n^2$ different lines each containing at least two points of P.*

Proof. Let $L_{u,v}$ be the number of pairs of points of P which determine a line that goes through at least u, but at most v points of P. From (5.1) and basic counting arguments we know that $L_{u,v} \lesssim \frac{n^2v^2}{u^3} + \frac{nv^2}{u}$ (see Exercise 6.3). Fix a constant C, and consider $L_{C,N/C}$.

Then

$$\begin{aligned} L_{C,N/C} &\leq \sum_{i=0}^{\lfloor \log(N) \rfloor} L_{C2^i,C2^{i+1}} \\ &= \sum_{i=0}^{\lfloor \log(N) \rfloor} O\left(\frac{4N^2}{C2^i} + 4CN2^i\right) \\ &= O\left(\frac{N^2}{C}\sum_{i=0}^{\lfloor \log(N) \rfloor} 2^{-i} + NC\sum_{i=0}^{\lfloor \log(N) \rfloor} 2^i\right) \\ &= O\left(\frac{N^2}{C}\right). \end{aligned}$$

In other words, for some $C_o > 0$ we have $L_{C,N/C} \leq C_o\left(N^2/C\right)$. Thus for the appropriate choice of C, at least half of the pairs of points determine a line through fewer than C, or at least N/C points. And consequently, at least a fourth of the pairs go through fewer than C points, or a fourth go through at least N/C points. In either case we are done. □

Consider a set, P, of n points and let $\mathcal{L}$ denote the set of lines passing through at least two points of P. An averaging argument (see Exercise 6.1) applied to Lemma 6.1 implies that there exists an absolute constant, c_o, such that at least $c_o n$ points of P are incident to at least $c_o n$ lines of $\mathcal{L}$. Then let B be the set of such points, and take some arbitrary point $a \in B$.

Draw in the lines through a that go through points of P. There must be at least $c_o n$ such lines. Choose one point other than a on each of these lines and draw in the circles around a that hit those chosen points (deleting those capturing fewer than 3 points). On each of these circles, break the points in triples, possibly deleting as many as 2 from each. We still have $\gtrsim n$ points left by our hypotheses (check!).

We call a triple "bad" if all three bisectors formed from its points go through at least k points. And we call the initial point a from B "bad" if at least half of its triples are bad. We would like to choose k such that at least half the points of B are bad. Clearly, the smaller k is the "easier" it is to get k-rich lines and thus more bad points.

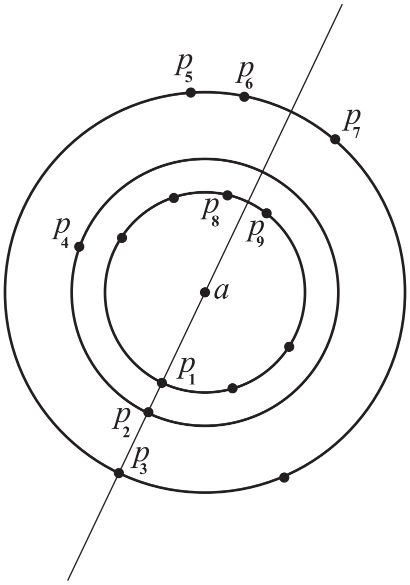


Figure 6.1. The point a is in B. Suppose we chose p_1, then p_2 and p_3 could not be chosen for a to contribute to the circles. The circle containing p_4 will be deleted. The points p_5, p_6, and p_7 form a triple. The point pairs (p_6, p_7) and (p_8, p_9) share a bisector with three points on it.

However, it will become clear that we would like k as large as possible. You will show in Exercise 6.2 that we may take $k = \frac{c_2 n^2}{t^2}$.

Then, if we can get the following upper and lower bounds on the number of incidences, $I(L_k, P)$, of k-rich lines and bad points, we will be done:

$$n^2/t^{2/3} \lesssim I(L_k, P) \lesssim t^4/n^2.$$

Finding an upper bound on $I(L_k, P)$ is straight forward. We simply apply (5.1) to find a bound on the number of k-rich lines, and then use Theorem 4.6 to get that $I(L_k, P) \lesssim n^2/k^2$. Getting a lower bound on the quantity $I(L_k, p)$ in terms of n and t is somewhat harder. The following Lemma is the key to the whole proof.

2. Arithmetic enters the picture

Lemma 6.2. *Let T be a set of N triples, (a_i, b_i, c_i), of distinct real numbers such that $a_i < b_i < c_i$ for $i = 1, \ldots, N$, and $c_i < a_{i+1}$ for all but at most $t-1$ of the i. Let $W = \{\frac{a_i+b_i}{2}, \frac{a_i+c_i}{2}, \frac{b_i+c_i}{2} : i = 1, \ldots, N\}$. Then $|W| \gtrsim \frac{N}{t^{2/3}}$.*

Proof. Let the range of a triple, $(a, b, c) \in T$, be defined as the interval $[a, c]$. By assumption, the sequence $(a_1, b_1, c_1, a_2, b_2, c_2, \ldots, a_N, b_N, c_N)$ be partitioned into at most t contiguous monotone increasing subsequences. Partition the real axis into $N/(2t)$ open intervals so that each interval fully contains the ranges of t triples. These intervals are constructed from left to right. Let x denote the right endpoint of the rightmost interval constructed so far. Discard the at most t triples whose ranges contain x, and move to the right until you reach a point y that lies to the right of exactly t new ranges. We add (x, y) as a new open interval, and continue in this manner until all triples are processed.

Let s be one of the open intervals defined in the previous paragraph. Let

$$S := \bigcup_{j:\{a_j,b_j,c_j\}\in s} \left\{\frac{a_j+b_j}{2}, \frac{a_j+c_j}{2}, \frac{b_j+c_j}{2}\right\}.$$

Each triple in T whose range is fully contained in s contributes three elements to $W \cap S$, and no two triples of T contribute the same triple to $W \cap S$. For every three elements in W that were contributed by elements in s, there is exactly one unique triple. Since there are $\binom{|W\cap S|}{3} \approx |W \cap S|^3$ ways to choose unique triples, $|W \cap S|^3 \gtrsim t$, the number of triples in the interval. It follows that $|W \cap S| \gtrsim t^{\frac{1}{3}}$, since otherwise the number of distinct triples of its elements would be smaller than t. If s' is another interval, with corresponding set S' contributing elements to W, notice that $S \cap S' = \emptyset$. Since the number of intervals, processed like s, is $N/(2t)$, the conclusion of the Lemma follows by the multiplication principle. □

For each point, $p \neq a$, in a bad triple, map p to the orientation of the ray $\overrightarrow{ap}$. By construction, this map is an injection, and W corresponds to k-rich lines. Therefore the number of k-rich lines incident to a is $\gtrsim n/t^{2/3}$. And since a was an arbitrary element of B, we get that $I(L_k, P) \gtrsim n^2/t^{2/3}$.

Recall that Exercise 6.2 shows that if we take $k = \frac{c_2 n^2}{t^2}$, then half of the points of P are "bad". Now we just write everything that we know together on one line.

$$\frac{t^4}{n^2} \approx \frac{n^2}{k^2} \gtrsim I(L_k, P) \gtrsim \frac{n^2}{t^{\frac{2}{3}}}.$$

A little bit of pencil pushing shows that this implies the desired bound. See [**38**] for the details, and some more specific hints on some of the Exercises.

Theorem 6.3 (Solymosi-Tóth [**38**])**.** *Let P be a set of n points in the plane. Then*

$$\#\Delta(P) \gtrsim n^{\frac{6}{7}}.$$

Exercises

Exercise 6.1. Write up the details of the averaging argument which tells us that "many" points go through "many" lines of $\mathcal{L}$. *Hint:* recall that we may assume that $t = o(n)$.

Exercise 6.2. Work out the details showing that we may take $k = \frac{c_2 n^2}{t^2}$ and at least $c_o n/2$ points of B will still be bad. Do this by constructing a multigraph, G, out of the points that are part of the triples as in the proof of Theorem 5.2. Find a way to draw g good edges for each point, where g is the number of good points. Next, apply the result of Theorem 4.8. Be sure to take into account the possiblity that $e < 5mv$.

Exercise 6.3. Check that (5.1) and basic counting arguments give us that $L_{u,v} \lesssim \frac{n^2 v^2}{u^3} + \frac{nv^2}{u}$.

Exercise 6.4. Find the constants C and C_o in the proof of Theorem 6.1, and write up the details of why we are done in the case where at least a fourth of the pairs go through at least N/C points of P.

Chapter 7

Beyond $n^{\frac{6}{7}}$

If you recall, the gain made from $n^{\frac{2}{3}}$ to $n^{\frac{4}{5}}$ came from considering the bisectors that were incident to many points. These bisectors came, of course, from pairs of points. Then, the gain from $n^{\frac{4}{5}}$ to $n^{\frac{6}{7}}$ came when we considered bisectors associated with triples of points. As you may imagine, more improvements have come from creatively considering quadruples of points, etc... Following this line of reasoning leads to many interesting questions and ideas. This Chapter will outline some of these, and hopefully convince you to keep exploring.

1. Sums and entries

We now introduce some new notation for the types of statements that we will be concerning ourselves with here. If k is a positive integer, call let α be the *strength* of a statement, $SE(k, \alpha)$, which we will describe below.

Definition 7.1. Consider an $M \times k$ matrix, A, with distinct entries. Let S be the set of all pairwise sums of entries of A in the same row. That is,

$$S := \{a_{ij} + a_{il} : j \neq l\}.$$

Define $SE(k, \alpha)$ to be the assertion that:

$$M << \#(S)^{\alpha}$$

.

The search for which values of k and α make the statement $SE(k, \alpha)$ true is called the *sums and entries problem.*

Exercise 7.1. Explain, in your own words, how the statement $SE(3, 3)$ is equivalent to Lemma 6.2.

Exercise 7.2. Trace through the reasoning in the previous Chapter, and show that if we replace portion involving Lemma 6.2 with a statement $SE(k_0, \alpha_0)$, that we are guaranteed to have $n^{\frac{4}{5-\frac{1}{\alpha_0}}}$ distinct distances. *Hint:* Show that in the Erdős distance problem setting, $SE(k_0, \alpha_0)$ implies that there are more than $\frac{n}{t^{1-\frac{1}{\alpha_0}}}$ different sums.

2. Tardos' elementary argument

Now that we have established a relationship between the Erdős distance problem and the sums and entries problem, by way of the previous two exercises, we will use ideas in the latter to improve results in the former. The following comes from [**46**].

Theorem 7.1. *Given an $M \times 5$ matrix, A, with distinct entries. Let S be the set of all pairwise sums of entries of A in the same row. That is,*

$$S := \{a_{ij} + a_{il} : j \neq l\}.$$

Then $M << \#(S)^{\frac{11}{4}}$. In other words, $SE(5, \frac{11}{4})$.

Proof. Let the size of S be n. We will call a number x a *heavy number*, with *weight* $\frac{1}{4}$, if it can be written as a difference of two elements of S in at least $n^{\frac{1}{4}}$ different ways. This can also be called the number of *representations* of a number x. Notice that the differences between entries on the same row can be expressed as differences of sums, since $a_{il} - a_{im} = (a_{il} + a_{ij}) - (a_{im} + a_{ij})$.

We will similarly define a *heavy row* to be a row with a pair of entries whose difference is heavy. If a given row has such pair of entries, we will call it a *light row.* We will show that both the number of heavy rows, and the number of light rows can be bounded above by $n^{\frac{11}{4}}$. Notice that this is very similar to the ideas behind the proofs of Theorem 1.1, and the high multiplicity edge deletion part of the proof of Theorem 5.2.

Since there are only n^2 total representations of numbers as differences of elements of S, we can be assured that there are at most $n^{\frac{7}{4}}$ heavy numbers. Now, if we focus our attention on heavy rows, we can see that each heavy row has some entry a_{il}, such that $a_{il} - a_{im}$ is a heavy number, and $a_{il} + a_{im}$ is in S. If we average these two numbers, we get a_{il} back. That is,

$$\frac{(a_{il} - a_{im}) + (a_{il} + a_{im})}{2} = a_{il}.$$

There can be no more than on the order of n averages for each heavy number, so in total, there are no more than $n^{\frac{11}{4}}$ of these types of averages. Since we know that entries in our matrix A are distinct, we can have no more than $n^{\frac{11}{4}}$ heavy rows.

We will now attempt to bound the number of light rows. Define the number $s_{lm}(i)$ to be the sum of the lth and mth entries in the ith row, that is,

$$s_{lm}(i) = a_{il} + a_{im}.$$

Clearly every such $s_{lm}(i) \in S$. Notice that there are no more than n^2 possible values for the pair $(s_{12}(i), s_{13}(i))$, where i indexes only light rows.

Since

$$s_{12}(i) - s_{13}(i) = a_{i2} - a_{i3} = s_{24}(i) - s_{34}(i),$$

and there are at most $n^{\frac{1}{4}}$ ways to represent $s_{12}(i)$ and $s_{13}(i)$, there are only $n^{\frac{9}{4}}$ possibe values of the quadruple

$$(s_{12}(i), s_{13}(i), s_{24}(i), s_{34}(i)).$$

If you iterate this argument two more times, by checking the existing quadruples against pairs $(s_{25}(i), s_{35}(i))$, you will get that there are, again, only $n^{\frac{1}{4}}$ different such pairs possible. So there are no more than $n^{\frac{10}{4}}$ different sextuples of the form

$$(s_{12}(i), s_{13}(i), s_{24}(i), s_{34}(i), s_{25}(i), s_{35}(i)).$$

If you continue again in this manner, you will get that there are at most $n^{\frac{11}{4}}$ different octuples of the form

$$(s_{12}(i), s_{13}(i), s_{24}(i), s_{34}(i), s_{25}(i), s_{35}(i), s_{15}(i), s_{45}(i)).$$

At last, we notice that

$$2a_{i2} = s_{24}(i) + s_{25}(i) - s_{45}(i).$$

This means that for each light row, the entry a_{i2} is completely determined by each possible octuple. So, since there are no more than $n^{\frac{11}{4}}$ distinct octuples, there can be no more than $n^{\frac{11}{4}}$ distinct light rows. This completes the proof of the Theorem. □

If we apply this new value of $\alpha = \frac{11}{4}$ as indicated in Exercise 7.2, we will get the following result, which slightly outdoes Theorem 6.3.

Theorem 7.2 (Tardos [**46**]). *Let P be a set of n points in the plane. Then*

$$\#\Delta(P) \gtrsim n^{\frac{44}{51}}.$$

Exercise 7.3. Describe how the idea behind rich bisectors are similar to the idea behind heavy numbers. In what ways are these ideas different? This is a very important point.

Exercise 7.4. Why did we stop at octuples?

3. Elementary Katz-Tardos method

We just showed $SE(5, \frac{11}{4})$, but even with $k = 5$, this estimate is not optimal. Nets Katz and Gábor Tardos were able to raise the bar and show that $SE(5, \frac{19}{7})$ was also true. Here we will only indicate the main ideas of the argument to keep from getting too bogged down with details. Through the exposition, we will allow the reader to complete the details via the included Exercises. We will show that for any $\epsilon > 0$, however small, $SE(5, \frac{19}{7} + \epsilon)$.

Exercise 7.5. Show that $SE(k, \alpha + \epsilon)$ implies $SE(k, \alpha)$. The fact that the matrices take real values has very little to do with what is really going on here. Try replacing the real number entries by vectors and see that $SE(k, \alpha)$ is still true for the same pairs of k and α as for the setting that we have studied so far. If we consider two matrices, A

and B which satisfy the hypotheses, then their *tensor product* matrix, C, will also satisfy these hypotheses.

$$c_{il,jm} = a_{ij} \cdot b_{lm}.$$

After verifying this, compare the resulting exponents.

We will try to show that $SE(5, \alpha)$, for every $\alpha > \frac{19}{7}$. First, we need to consider heavy numbers of weight $3 - \alpha$. That is, numbers which can be represented as a difference of elements in S in more than $n^{3-\alpha}$ ways. This will give us the desired number of heavy rows, as before. Again, we will present an argument showing an identical bound for light rows.

The main new idea here is that we will now begin considering phenomena involving pairs of light rows, as opposed to just one light row at a time. We want to consider the pairs of rows, (i, i'), such that

$$s_{12}(i) = s_{12}(i'), s_{23}(i) = s_{23}(i'), s_{34}(i) = s_{34}(i'), s_{45}(i) = s_{45}(i').$$

We will call V the set of such pairs of light rows. The goal here is to get a handle on how large V can be. Since there are n^2 choices for $(s_{12}(i), s_{23}(i))$, and the rows in question are light, there are $n^{8-2\alpha}$ choices of quadruples of the form

$$(s_{12}(i), s_{23}(i), s_{34}(i), s_{45}(i)).$$

Exercise 7.6. Use the Cauchy-Schwarz inequality to show that $\#V \geq n^{4\alpha-8}$. *Hint:* The quantity $\#V$ is a sum of squares, and the inequality is sharp when each choice of quadruples $(s_{12}(i), s_{23}(i), s_{34}(i), s_{45}(i))$ occur for equally many light rows i.

Now, if the pair (i, i') is in V, then we are guaranteed that

$$a_{i1} - a_{i3} = a_{i'1} - a_{i'3},$$

and

$$a_{i3} - a_{i5} = a_{i'3} - a_{i'5}.$$

Now we define a funciton ν on V as

$$\nu(i, i') = s_{13}(i) + s_{35}(i'). \tag{7.2}$$

We can also observe that

$$\nu(i,i') = s_{13}(i') + s_{35}(i), \tag{7.3}$$

and

$$\nu(i,i') = 2a_{i3} + s_{35}(i'). \tag{7.4}$$

These three equivalences of $\nu(i,i')$ will take us through to our conclusion.

Exercise 7.7. Show that the numbers $s_{15}(i')$ and $\nu(i,i')$ uniquely specify the pair (i,i'). *Hint:* Using (7.4), we can uniquely determine a_{i3} which, in turn, uniquely specifies i. If we use the definition of V and the fact that we know $s_{15}(i)$, we can find i' the same way.

Using Exercise 7.7, we know that there are n elements of V on which ν takes on some specific value ν_0. Since we know two methods of finding different elements of V with the same value of ν. One way is to find different pairs (i,i') with the same values of $s_{13}(i)$ and $s_{35}(i')$, and the other way is to find such pairs with the same values of $s_{13}(i')$ and $s_{35}(i)$.

Now we will offer a heuristic argument, which you will clean up using Exercise 7.9. Since there are $N^{4\alpha-8}$ elements of V and n^2 possible values of the function

$$B(i,i') = (s_{13}(i), s_{35}(i')),$$

the typical level set [1] of B should have about $n^{4\alpha-10}$ elements. Similarly, define the function

$$C(i,i') = (s_{13}(i'), s_{35}(i)).$$

For each element in some level set of B, list all of the elements of V that are in the same level set of C. That is, pick some value of the function B, and find all of the pairs that give that value. For each such pair, find all of the other pairs that are in the first pair's level set of C. Listing pairs in this way will give $n^{8\alpha-20}$ elements total. Of course, this will be overcounting, as we have listed each element as many times as the size of the joint level set of the function (B,C),

[1] A level set in this sense is a set of pairs of light rows, (i,i'), such that $B(i,i')$ is equal across all pairs in the level set. You can think of it as the set of "points" that all have some value of B.

which will be the set of pairs which return equivalent values for both B and C.

Exercise 7.8. Show that specifying $B(i, i')$, $C(i, i')$, and i specifies i'.

So the size of the joint level set of B and C should be the same as the number of light rows, i with $s_{13}(i)$ and $s_{35}(i)$. Since there are n^{α} rows, if all of the joint level sets of $(s_{13}(i), s_{35}(i))$ are equally sized, the level sets should have size $n^{\alpha-2}$. Thus, we should be able to find a level set of ν that has size $n^{7\alpha-18}$. However, Exercise 7.7 tells us that no level set of ν can have size greater than n. So by comparing upper and lower boounds on the size of any level set, $\alpha \leq \frac{19}{7}$.

Exercise 7.9. Let f be a function from a finite set X to the interval $[1, N]$. Show that there is a subset, $Y \subset X$, and a number $\rho \in [1, N]$ for which

$$|Y| \geq \frac{|X|}{\log N},$$

and such that for every $y \in Y$, we have

$$\rho \leq f(y) \leq 2\rho.$$

This is called the *dyadic pigeonhole principle*, and we have already used it in proving Theorem 5.2. Do you remember where?

Exercise 7.10. Apply Exercise 7.9 repeatedly, and then use Exercise 7.5 to complete a rigorous proof of the main result of this section, $SE(5, \frac{19}{7})$ and as before, show that it proves Theorem 7.3.

Theorem 7.3 (Katz-Tardos [**25**]). *Let P be a set of n points in the plane. Then*

$$\#\Delta(P) \gtrsim n^{\frac{19}{22}}.$$

To sum up what we have accomplished so far, $\frac{6}{7} \approx .857142$, $\frac{44}{51} \approx .862745$, and $\frac{19}{22} \approx .863636$. The world record as of this writing is also due to Katz and Tardos, $\frac{48-14e}{55-16e} \approx .864137$. The next section will introduce an example by Imre Ruzsa, [**37**], which seems to limit the possible development of approaches of this style.

4. Ruzsa's construction

Although the sums and entries problem has borne much fruit, it appears as though it has a distinct upper bound to just how close it can get us to the full Erdős conjecture. As Exercise 7.2 indicates, if we could show that there was some sequence of values of k for which α_k approached 1, we would have a positive solution to the Erdős distance problem. However, the following construction makes it look as though this train of reasoning will derail before solving the whole problem.

We will start by writing down a long list of vectors whose entries are 1 or -1, whose pairwise dot products are small negative numbers. More precisely, we will construct k vectors, and show that the values of α_k associated with each such k will approach 2, which would lead us to believe that there is a limit of $\frac{8}{9}$ for the Erdős exponent.

Let k be even, and define the vectors $a_1, \ldots, a_k$ to be of dimension

$$m = \binom{k}{\frac{k}{2}}.$$

We will identify coordinates with subsets of size $\frac{k}{2}$. If D is such a subset, then the Dth component of a_i will be written as a_{iD}, and will be equal to 1 if $i \in D$, and -1 otherwise. Now we will appeal to the fact that given two distinct elements of $\{1, \ldots, k\}$, and a random subset, D, of size $\frac{k}{2}$, the probability that both elements are in D is a little less than $\frac{1}{4}$, as is the probability that both elements are not in D.

Exercise 7.11. With $a_1, \ldots, a_k$ and m as above, show that if i and j are distinct, then the dot product of a_i and a_j, denoted $a_i \cdot a_j$ is $\frac{-m}{k-1}$.

Exercise 7.12. If k is odd, construct vectors $a_1, \ldots, a_k$ of 1's and -1's, of dimension m, such that for any distinct i and j, $a_i \cdot a_j$ is $\frac{-m}{k}$. *Hint:* Think of what happened in the previous Exercise.

After constructing vectors as in the previous two Exercises, we will construct counterexamples to show that the claim that $SE(k, \alpha)$ must be false for some values of α. We will work with even k, but odd k works very similarly.

Construct an $n \times k$ matrix, A, of vectors of dimension N, with

$$n = \binom{N}{m},$$

where m is the dimension of the vectors we constructed before, and N is chosen to be as large as we want. Let $e_1, \ldots, e_N$ stand for the *canonical basis* for an N dimensional vector space, or the set of vectors where each e_j has a zero in each coordinate, except for a 1 in the jth coordinate. We will identify rows of the matrix we are constructing with choices $t_1 < \cdots < t_m$ of m coordinates in our N dimensional vector space. We will let the entries of A be the images of the a_j's that we constructed earlier, but in the m coordinate positions that we have chosen. That is, if we call σ our list of $t_1 < \cdots < t_m$, then

$$A_{\sigma j} = \sum_{l=1}^{m} a_{jl} e_{t_l}.$$

All of the entries of A are distinct, as the row determines the positions of the non-zero entries of the vector, and the column determines what the entries are. The sums are vectors whose non-zero entries are either 2 or -2, and have a relatively small number of such entries, per our dot product condition. In fact, they will have exactly m' non-zero entries, where

$$m' = \frac{(k-2)m}{2(k-1)}.$$

We will choose N to be so large as to ignore constants which depend on m. The number of sums will be bounded by

$$2^{m'} \binom{N}{m'} \approx N^{m'},$$

but the number of rows is on the order of N^m, which shows that $SE(k, \alpha)$ cannot be true if α is grater than $\frac{k-2}{2k-1}$, which approaches 2 as k grows large.

Exercise 7.13. Get an analogous result for odd values of k.

This marks the end of the first part the book. From here on out, the flavor of the will change slightly. The difficulty will increase a little, and the settings will vary even more drastically. We will start exploring other types of problems that are inspired by or related to the study of the main Erdős distance problem. This is quite important to see, as regardless of the inherent beauty of any problem, without some context or relevance, it lacks some of its luster. In continuing through the next few chapters, you should also try to pick up on how these problems are related. Try to find salient features that are present in some or all of the different settings. This way, you can see how mathematicians use ideas from the study of one problem in the study of another.

Chapter 8

Information theory

In this Chapter, we will introduce a few of the ideas of information theory. The theory is beautiful in its own right, but our main motivation is, of course, its relationship with the Erdős distance problem. The main thrust will be to elucidate the information theoretic interpretation of the sums and entries problem, which, as we have seen, is related directly to the Erdős distance problem.

1. What is this information of which you speak?

Information theory is a branch of probability theory, which concerns itself largely with the study of random variables. One way of thinking about random variables is that they are a model for our knowledge of the universe. We might not know the precise outcome of some particular event, say exactly where a ball will land if we throw it straight up in the air, but if we throw it and observe where it lands, we will have a clearer idea with about where it may land for subsequent tosses. Information theory studies this very phenomenon, the amount of information learned by *collapsing* a random variable, or performing an experiment and observing the outcome.

If A is a random variable with possible outcomes $a_1, \dots, a_m$, and associated probabilities $p_1, \dots, p_m$, respectively, then we can define the *entropy*, $H(A)$, associated with the random variable A by

$$H(A) = -\sum_{i=1}^{m} p_i \log p_i.$$

Although this definition might look puzzling at first, try to think about $H(A)$ as some quantity of information. We will now offer several explanations of where this comes from, to hopefully clear up the intuition before we use it.

Computers operate on *bits* of information, which are typically thought of as 1's and 0's. We can think of these bits as the amount of information needed to distinguish between two possibilities which are equally likely. Clearly we can distinguish between 2^m equally likely events with m bits. This should somewhat justify the need for about $-\log p$ bits of information to distinguish among $\frac{1}{p}$ equally likely events.[1] If all of our considered events were equally likely, then our definition of $H(A)$ would be relatively secure. However, not every event is equally likely, so we have some explaining to do.

Exercise 8.1. What probability would you assign to the event that the world ends today? How suprised would you be if you found out, through some reliable channel, that the world was not ending today? How suprised would you be if you found out, with just as much reliability, that it was ending today?

Try to think of the $-\log p$ as the amount of suprise if an event with probability p actually occurs. If this still seems murky, you are not alone. For this reason, we have two more alternate explanations. One is from a mathematical point of view, available in [**26**], and the other is from physics, [**23**]. We will summarize their explanations below.

The mathematical explanation assumes that each proability is a rational number with denominator n. Collapsing the random variable, A, is part of a two step process. First, we need to observe which outcome has occurred. Assume that it is a_j. Next we need to choose between np_j of the possible equally likely outcomes that are part of the event a_j. (Why are there np_j equally likely outcomes corresponding to a_j again?) Now, the total information gained by choosing from

[1] In this instance, the log has base 2.

among n equally likely outcomes is $\log n$. So the expected information gained gained by performing the second task is

$$\sum_{j=1}^{m} p_j \log(np_j).$$

The leftover information is then $H(A)$.

The explanation from physics assumes that the experiment can be performed independently many times. So we repeat the experiment N times, for some large N. We will expect a_j to have about $p_j N$ outcomes which correspond to the event a_j. The number of ways to arrange these outcomes is

$$\frac{N!}{(p_1 N!) \dots (p_m N!)}.$$

Since all possible orderings are equally likely, we get

$$NH(A) = \log\left(\frac{N!}{(p_1 N!) \dots (p_m N!)}\right).$$

If we appeal to Stirling's formula, which says that $\log(N!) = N \log N - N + L$, where $L \approx \log N$, we get the formula for $H(A)$ that we gave earlier.

Exercise 8.2. Explain the connections between these heuristic explanations.

2. More information never hurts

Often times, when teaching a calculus class, we have found that students want us to work homework problems more than teach them general theory of calculus. We try to convince the students that the homework problems should be easier if the students pay attention to the theory lectures first. In other words, the basic idea we want to address here is that more information never hurts. We will exploit the fact that the function $x \log x$ is convex when x is positive.

Proposition 8.1. *The information, $H(A)$, is maximized when each $p_j = \frac{1}{m}$.*

Proof. Recall the definition of $H(A)$,

$$H(A) = -\sum_{i=1}^{m} p_i \log p_i.$$

If we let $f(x) = x \log x$, we can write this as

$$H(A) = -\sum_{i=1}^{m} f(p_i) = -m \sum_{i=1}^{m} \frac{1}{m} f(p_i).$$

If we now appeal to Jensen's inequality, with $\frac{1}{m}$ taken as the probability distribution, we get

$$H(A) \leq -mf \left(\sum_{i=1}^{m} \frac{1}{m} p_i \right) = -mf \left(\frac{1}{m} \right) = \log m,$$

as promised. □

This is all well and good, but what happens when there are two random variables? We will keep A as before, and let B be a new random variable with possible outcomes $b_1, \ldots, b_m$, with associated probabilities $q_1, \ldots, q_n$, respectively. Suppose that A represents our theoretical knowledge of calculus, and B represents our ability to solve particular homework problems. Our goal is to show that no more information is required in resolving B if we have resolved A or if we have not.

Let (A, B) stand for the random variable which is the joint outcome of A and B. This has the possible outcomes (a_i, b_j), and probabilities $p_i q_j$. We will define the entropy of this random variable to be the *joint entropy* of A and B, and write it $H(A, B)$. If we fix an index i, the numbers $q_{ij} := p_i q_j$ will form a probability distribution. That is to say, if we know a particular value of a_j is the outcome of A, then we might gain more information on the outcome of B. These will be referred to as the probabilities of b_j *conditional* on a_i, written

$$\mathbb{P}(B = b_j | A = a_i) = q_{ij}.$$

Given the setting above, we now introduce Bayes' law,

$$\sum_{i=1}^{m} p_i q_{ij} = q_j,$$

which will be instrumental in the proof of the following Theorem.

Theorem 8.2. $H(A,B) \le H(A) + H(B)$.

We can sum up the above statement as, "Extra information doesn't hurt." If I roll two dice and look at only one of them, it doesn't hurt my chances of guessing the other one.

Exercise 8.3. In this context, we say that the random variables A and B are independent if $q_{ij} = q_j$ for all i. Verify that in this case, we have the equality

$$H(A,B) = H(A) + H(B).$$

Proof.

$$H(A,B) = -\sum_{i=1}^{m}\sum_{j=1}^{m} p_i q_{ij}(\log p_i + \log q_{ij}) = H(A) - \sum_{i=1}^{m}\sum_{j=1}^{m} p_i q_{ij} \log q_{ij}.$$

We will define the second term in the right hand side as $H(B|A)$, the entropy of B *conditional* on A. So, to prove the statement, we need only show that $H(B|A) \le H(B)$. This statement, in our situation, can be translated as, "Homework problems are, no harder, and probably easier if you know theory." Continuing, in terms of f, where $f(x) = x\log(x)$, and using Jensen's inequality followed by Bayes' law.

$$H(B|A) = -\sum_{i=1}^{m}\sum_{j=1}^{m} p_i f(q_{ij}) \le -\sum_{j=1}^{m} f(p_i q_{ij}) = -\sum_{j=1}^{m} f(q_j) = H(B).$$

□

Basically, $H(B|A)$ is the expected information of B conditioned on a random value of A, that is

$$H(B|A) = \sum_{i=1}^{m} p_i H(B|A = a_i).$$

The point is that the conditional information $H(B|A)$ is not really the information of any random variable in particular, but it is a linear combination of the informations of random variables.

We say that a random variable X *determines* a random variable Z if the outcome of X determines the outcome of Z.

Exercise 8.4. Show that if the random variable X determines the random variable Z that

$$H(Z) \leq H(X).$$

Exercise 8.5. Show that if the random variable X determines the random variable Z, that

$$H(X, Z) = H(X).$$

Hint: Write out the definitions, and think about what the various probabilities will be if one random variable determines another.

We now introduce the *submodularity principle*.

Theorem 8.3. *Let A, B, X, and Y be random variables such that each of X and Y determine B, and that X and Y jointly determine A, then*

$$H(A) + H(B) \leq H(X) + H(Y).$$

Proof. Recalling the definition of conditional information, Exercise 8.5, and the fact that X and Y determine B, if we subtract $2H(B)$ from each side from the claim, we get

$$H(A|B) \leq H(X|B) + H(Y|B),$$

which is what we will prove.

Since X and Y jointly determine A, we can use Exercise **??** that $H(A) \leq H(X, Y)$, which implies that $H(A|B) \leq H(X, Y|B)$. So now we are reduced to showing that

$$H(X, Y|B) \leq H(X|B) + H(Y|B).$$

If we employ Theorem 8.2 for each outcome, then we get

$$H(X, Y|B = b_i) \leq H(X|B = b_i) + H(Y|B = b_i),$$

and we need only take expected values on both sides to finish. □

Exercise 8.6. Give a necessary and sufficient condition for Theorem 8.3 to be sharp. *Hint:* Think about how this resembles the Cauchy-Schwarz inequality, and what the sharp case was there.

3. Application to the sums and entries problem

After all of that development of information theory, we will sketch out the ideas used to make improvements to the sums and entries problem, and subsequently, to the Erdős distance problem. The full arguments are available in [**25**] and [**46**], and after you finish this portion of the book, you will be fully equipped to tackle them in full detail.

First, we will formulate the sums and entries question as an information theoretic question. Let A be an $N \times s$ matrix with distinct entries. We will define $S(A)$ as before, to be the set of sums of entries of A which are in the same row. We are looking for lower bounds on $M = \#S(A)$ of the form $N \leq M^{\alpha}$, which we can rewrite as

$$\log N \leq \alpha M.$$

We will view this as an inequality between quantities of information. Let R be a random variable whose value is a row of A. Each row can be chosen with probability $\frac{1}{N}$. We will define a class of functions on R, the patterns p_{UV}, with U and V as subsets of $\{1, \ldots, s\}$. We will define $p_{UV}(R)$ to be a set consisting of all of the sums, $R_i + R_j$ with $i \in U$ and $j \in V$, and all of the differences $R_i - R_j$ for either $i, j \in U$ or $i, j \in V$. So $p_{UV}(R)$ is also a random variable, and we denote by $H(U, V)$ the information

$$H(U, V) = H(p_{UV}(R)).$$

Certain facts involving the $H(U, V)$'s follow immediately from the basic priciples of information theory.

(i) $H(U, V) = H(V, U)$.

(ii) $H(U, V) \leq H(U', V')$ if $U \subset U'$ and $V \subset V'$.

(iii) $H(U, V) = 0$ if U is empty and $\#V = 1$.

(iv) $H(U, V) \leq \log |S(A)|$ if $U \neq V$ and $\#U = \#V = 1$.

(v) $H(U, V) = \log N$ if $U \cap V$ is not empty and $\#(U \cup V) > 1$.

(vi) $H(U \cup U', V \cup V') + H(U \cap U', V \cap V') \leq H(U, V) + H(U', V')$ if $(U \cap U') \cup (V \cap V')$ is not empty.

Exercise 8.7. Prove statements (i)-(vi). *Hint:* (iv) uses random variables with uniform distributions that have the largest possible information, (v) uses the fact that entries are distinct, and (vi) uses the submodularity principle.

Now, the set $\{1,\ldots,s\}$ has 2^s subsets, which might seem like a lot. We would like to summarize these prior statements by averaging them somehow. For $i,j \geq 0$ and $1 \leq i+j \leq s$, we will define the normalized information average, $H_{i,j}$ by

$$H_{i,j} = 1 - \frac{1}{\binom{s}{i}\binom{s-i}{j}\log N}\sum_{U,V} H(U,V),$$

where the sum is over disjoint subsets U and V, for which there are clearly $\binom{s}{i}\binom{s-i}{j}$ choices. We then get the following:

(vii) $H_{i,j} = Hj, i$.

(viii) $H_{i,j} \leq H_{i+1,j}$ if $i+j \leq s-1$.

(ix) $H_{0,1} = 1$.

(x) $H_{1,1} \geq 1 - \frac{\log \#S(A)}{\log N}$.

(xi) $H_{i-1,j} + H_{i+1,j} \geq 2H_{i,j}$ if $i \geq 1$ and $2 \leq i+j \leq s-1$.

(xii) $H_{i,j} \geq H_{i+1,j} + H_{i,j+1}$ if $i+j \leq s-1$.

Exercise 8.8. Prove (vii)-(xii) using (i)-(vi).

Hopefully that was not too hard. Now we are ready to say something nontrivial about the sums and entries problem. Actually, using just these facts, (vii)-(xii), it is possible to use *linear programming*[2] to find bounds for $H_{1,1}$, which gives $SE(k,\alpha_k)$ with α_k approaching e, the base of the natural logarithm.

Exercise 8.9. Prove that $SE(5, \frac{11}{4})$ is true using only the facts (vii)-(xii). Then get a smaller α value with $k = 7$.

If we can deduce one more fact, which is very similar to the argument given for the validity of the statement $SE(5, \frac{19}{5})$:

(xiii) $5H_{1,1} - H_{2,1} + 2H_{3,0} \leq 3$.

[2] Linear programming typically refers to a set of linear constraints or inequalities, under which some quantity is to be maximiezed or minimized. A very simple example in the plane would be to find the the largest value of y subject to the constraints $y \leq 2x$ and $y \leq 10 - 3x$.

If you add this to your bag of tricks, and sprinkle in a bit of linear programming, you can show $SE(k, \alpha_k)$ is true for α_k approaching $\frac{24-7e}{10-3e}$.

Exercise 8.10. Prove the lucky inequality, (xiii). *Hint:* This is an adaptation of the proof of $SE(5, \frac{19}{7})$. Consider pairs of rows, (R, T) such that $p_{U\emptyset}(R) = p_{U\emptyset}(T)$ with $U = \{i, j, k\}$, a set of three elements, and assign the pairs (R, T) the following non-uniform probability distribution. Select R uniformly, and select T uniformly among those rows which satisfy our conditions with the given R. The advantage here is that $H((R, T)) = 2H(R) - H(p_{U\emptyset}(R))$ because we are in the sharp case of the submodularity principle. Then consider the function

$$\nu((R, T)) = R_i + R_k + 2T_j = R_i + R_j + T_j + T_k = R_k + R_j + T_j + T_i.$$

Use these three equalities and the submodularity principle to obtain the desired result.

Chapter 9

Dot products

The title of this book advertised distances, and now you are reading about dot products. Why? Well, at this point, you have seen the basic arguments that lead toward increasing lower bounds on the number of distinct distances determined by a large number of points in the plane. Now, the reason this chapter is here is not just so you can see how many distinct dot products are determined by a set of points in the plane. The main goal here is to illustrate how you can apply similar techniques in different settings. As you read through this chapter, try to pick out which key features of both problems would lead you to approach this problem as the distance problem.

1. Transeferring ideas

Given any $x, y \in \mathbb{R}^2$, we write their *dot product* as $x \cdot y$. If $x = (x_1, x_2)$ and $y = (y_1, y_2)$,

$$x \cdot y = x_1 y_1 + x_2 y_2.$$

There are other useful ways to think of the dot product, but this one will suffice for the arguments to follow.

Now, if we are given a set, P, of n points in the plane, define $\Pi(P)$ to be the set of all distinct dot products. As before, how many distinct dot products can we be sure to find? Before you go any further, try to work out how you could treat this question like the distance question. Specifically, what are the "circles" here?

Well, when we looked at distances determined by a single point, x, we noticed that distinct distances lay on distinct circles, all centered at x. So all the points of a given distance to x lie on the same circle. In this setting, given a point, x, what do all of the points that have the same dot product with respect to x lie on? In Exercise 9.1, you will show that they all lie on lines perpendicular to, l, the line between x and the origin. We call a line *radial* if it passes through the origin. So the line, l, mentioned before could be described as the radial line through x.

Now we are ready to apply ideas similar to those we used in Chapter 1, where we achieved $\sqrt{n}$ distinct distances. Recall that when we were given a set, P, of n points, we could pick a point in particular, x, and draw circles around it that covered the rest of the points. We found that either there were $\sqrt{n}$ circles around x, or there was a circle around x with $\sqrt{n}$ points on it. How could we do this for dot products?

First, let's pick a point out of P. Of course, we'll call it x. Now, we know that dot products with respect to x all lie on parallel lines. So let's count them. Suppose it takes t lines to cover our point set, P. Now, if $t \gtrsim \sqrt{n}$, we will have at least $\sqrt{n}$ distinct dot products with respect to x. What if t is significantly less than $\sqrt{n}$? By the pigeonhole principle, we know that one of the lines, l, will have at least $\frac{n}{t}$ points on it. Since we decided that $t < \sqrt{n}$, we can be assured that $\frac{n}{t} > \sqrt{n}$. So we now have a line with $\sqrt{n}$ points on it. Pick some point on l, call it y, that does not lie on the line through both x and the origin. Now, notice that covering the other points on l with another set of parallel lines, each perpendicular to the line between y and the origin, gives us $\sqrt{n}$ populated lines. Recall that each of these lines represents different dot products with respect to y. So either x or y will determine at least $\sqrt{n}$ distinct dot products. We have just shown the following to be true:

Theorem 9.1. *Let P be a set of n points in the plane. Then*

$$\#\Pi(P) \gtrsim n^{\frac{1}{2}}.$$

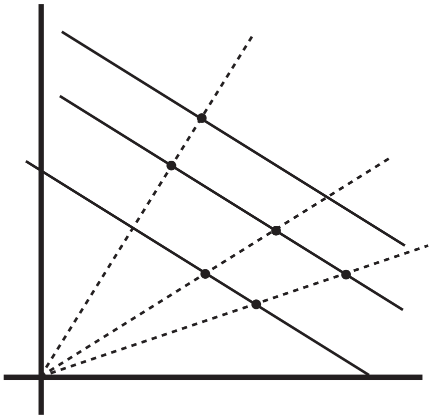


Figure 9.1. Drawing parallel lines that are perpendicular to the radial line through the upper two points.

Be sure to go back through the first proof of $\sqrt{n}$ for distances and the proof above for dot products and look for the subtle differences between the two.

2. Székely's method

If you look back to Chapter 5, and the ideas contained therein, you might be able to guess where this section is going. The last proof idea followed with very little change, and gave us identical results. Here we will see how to cope with differences between settings, and what results from that.

If we are given a set, P, of n points, the first thing we will do is construct a graph, G, similar to the one in the proof of Theorem 5.2. So, define the vertex set to be the point set P. Now we have to decide how to construct edges. In Székely's original argument, edges were drawn between points along circles. These circles were, of course, centered at points in our set. As before, what object in the dot product setting behaves similarly to circles? That would be the parallel lines perpendicular to the radial lines of points in our set. So after drawing these parallel lines, perpendicular to the radial line of

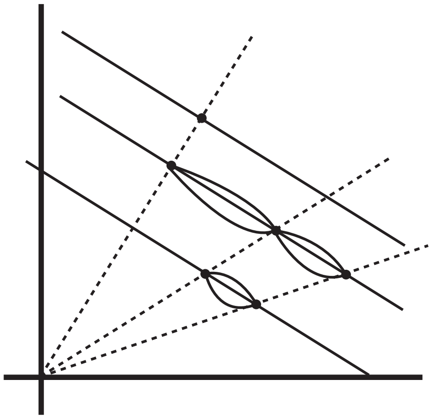


Figure 9.2. Drawing some edges associated with the leftmost radial line.

each point, for each point, we will draw edges between consecutive points along these lines.

Now, if there are several points along a radial line, these points will all define the same set of parallel lines. So we will have multiple edges between pairs of points along those lines.

Let t be defined as in the last proof. Now construct G by letting the points in P act as vertices. For each point, x, in P, draw an edge between pairs of consecutive points along the parallel lines which are perpendicuar to the radial line through x. Since we cover our point set with about n edges for each point in P, there must be about n^2 edges in G. So $e \approx n^2$. (We would already win if it took less than about n^2 edges to cover our point set. Why?)

If we consider a fixed radial line, l, the vertices of consecutive points along all of the parallel lines perpendicular to l will be connected by as many edges as there are points on l. We recall from the proof of Theorem 9.1 that there can be no more than t points along any line, much less a radial line. So no pair of vertices can be connected by more than t edges. So in our graph, the maximum edge multiplicity will be t.

We can apply the modified crossing number theorem, Theorem 4.8 to get:

$$\frac{n^6}{tn^2} \lesssim \frac{e^3}{mv^2} \lesssim cr(G).$$

Now we just need an upper bound for the crossing number. Note that a crossing between edges can only occur if a line perpendicular to one point's radial line crosses a line perpendicular to another point's radial line. Since each point has fewer than t such associated parallel lines, each pair of points can contribute at most t^2 crossings. There are about n^2 different pairs of points so the total number of crossings is definitely less than n^2t^2. So we can certainly bound the crossing number above by n^2t^2. Putting the upper and lower bounds for the crossing number together:

$$\frac{n^6}{tn^2} \lesssim cr(G) \lesssim n^2t^2.$$

So now we have shown the following theorem:

Theorem 9.2. *Let P be a set of n points in the plane. Then*

$$\#\Pi(P) \gtrsim n^{\frac{2}{3}}.$$

So we followed the idea in the proof for $n^{\frac{4}{5}}$ for distances, but we ended up with $n^{\frac{2}{3}}$. What was different? Of course, we never tried to lower the edge multiplicity. What happens if we try to? You will explore that in Exercise 9.4.

3. Special cases

In general, we had trouble reducing edge multiplicity. However, we can find some special classes of sets where we can do better than in the general case. Here we illustrate the idea of using techniques that could have limitations in some bad cases, and eliminate those cases. Below is an odd looking theorem. It has a strange and seemingly artificial condition about the number of points along a line through the origin. Soon enough though, we will see how we can use a theorem like this to prove some interesting corollaries.

Theorem 9.3. *Let $\#P = n$ and have no more than n^x points on any line through the origin. Then $\#\Pi(P) \gtrsim n^{1-\frac{x}{2}}$.*

Proof. Recall that in the graph theoretic proof of the dot product set result, we get

$$\frac{n^6}{mn^2} \lesssim \frac{e^3}{mv^2} \lesssim cr(G) \lesssim n^2t^2.$$

Now since no line through the origin has more than n^x points on it, no edge multiplicity is higher than n^x. So we can run the same argument with $m = n^x$, and get

$$t \gtrsim n^{1-\frac{x}{2}}.$$

□

Suppose that you have two sets of real numbers, A and B. We write *Cartesian product* of A and B as $A \times B$. It is defined as the set of all pairs of numbers, (a, b), where $a \in A$, and $b \in B$.

Corollary 9.4. *Let $P = A \times B$, where $A, B \subset \mathbb{R}$, and $\#P = n$. Let $min(\#A, \#B) = n^x$. Then $\#\Pi(P) \gtrsim n^{1-\frac{x}{2}}$.*

Cartesian product sets come up quite often in practice, but this is not the only kind of set that will obey the line condition in Theorem 9.3. There are plenty of times where we want to deal with sets that are sufficiently spread out in some sense. We introduce here a formal way to define a point set that is sufficiently spread out.

Definition 9.1. We call a set of size n, *well-distributed* if it has exactly one point inside each of an $n^{\frac{1}{2}}$ by $n^{\frac{1}{2}}$ lattice of squares with side length C, where C can be any specified positive constant.

Corollary 9.5. *Let P be well-distributed, and $\#P = n$. Then $\#\Pi(P) \gtrsim n^{\frac{3}{4}}$.*

Proof. Any line that passes through the set P may pass through no more than $2n^{\frac{1}{2}}$ squares. So at most, no line through the origin can pass through more than $cn^{\frac{1}{2}}$ points. So by Theorem 9.3,

$$\#\Pi(P) \gtrsim n^{1-\frac{1}{2}\cdot\frac{1}{2}} \gtrsim n^{\frac{3}{4}}.$$

□

The next definition is quite involved. It might help to think of a picture inside of a picture inside of a picture...

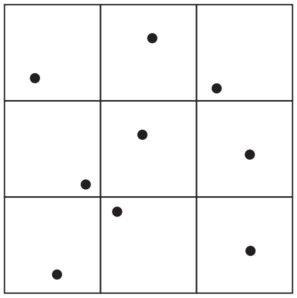


Figure 9.3. Example of a Well-Distributed set.

Definition 9.2. We call a set of size n, *2-iterated well-distributed* if it is comprised of $n^{\frac{1}{2}}$ translated well-distributed subsets, where each subset has constant C, and the subsets are each contained in one square of an $n^{\frac{1}{4}}$ by $n^{\frac{1}{4}}$ lattice of squares with side length $\max(C^2, C^{-2})$. Similarly, a set is *r-iterated well-distributed* if it is comprised of $n^{\frac{1}{r}}$ translated $(r-1)$-iterated well-distributed subsets, where each subset has constant $\max(C^r, C^{-r})$, and the subsets are each contained in one square of an $n^{\frac{1}{2r}}$ by $n^{\frac{1}{2r}}$ lattice of squares with side length $\max(C^r, C^{-r})$.

Note, by the above definitions, well-distributed is the same as 1-iterated well-distributed.

Corollary 9.6. *Let $\#P = n$, and let P be r-iterated well-distributed, where $r \leq log(n)$. Then $\#\Pi(P) \gtrsim n^{\frac{3}{4}}$.*

Proof. As in the proof of Corollary 9.5, the maximum number of large squares any line can pass through is at most $cn^{\frac{1}{2r}}$. Then, in

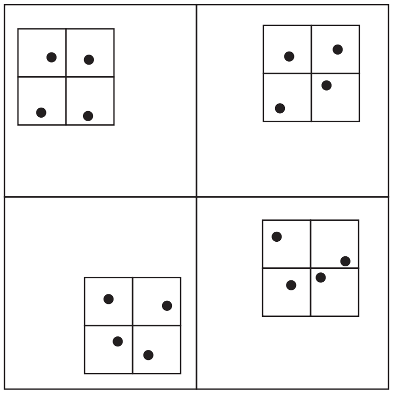


Figure 9.4. Example of a 2-Iterated Well-Distributed set.

each square, the most number of subsquares any line can pass through is at most $cn^{\frac{1}{2r}}$. This continues for r stages of iterations, so the total number of points any line can pass through is certainly at most $cn^{\frac{1}{2}}$. □

Even though these conditions are more natural looking, they still might not be quite what we need. Sometimes we will have to deal with a set that is, in some sense, almost in one of these classes. What do we mean by "almost" here? For example, if we have a set of n points, R, which is similar to a well-distributed set, but it has as many as ten points in each box. We can pick one point from each box, and call that point set R'. Now R' will be well-distributed, and it satisfies the conditions of Theorem 9.5. Since $\Pi(R') \subset \Pi(R)$, we can use this theorem to get the same exponent for R.

Exercises

Exercise 9.1. Show that given $s \in \mathbb{R}$, and a point $x = (x_1, x_2) \in \mathbb{R}^2 \backslash \{0\}$, all the points $y = (y_1, y_2) \in \mathbb{R}^2$ that satisfy the equation $x \cdot y = s$ lie on the line:

$$y_2 = \frac{s}{x_2} - \frac{x_1}{x_2} y_1.$$

What are the points that have constant dot product with the origin? Can any point have a non-zero dot product with the origin?

Exercise 9.2. In the proof of Theorem 9.1, why couldn't y lie on the line between x and the origin?

Exercise 9.3. Try to write a proof of Theorem 9.1 using the ideas in the second proof of Theorem 1.1.

Exercise 9.4. Try to emulate the edge counting scheme as in the proof of Theorem 5.2. What goes wrong?

Exercise 9.5. Given a large finite set, A, of real numbers, we define the set $AA + AA$ to be $\{ab + cd\}$, where $a, b, c, d \in A$. How big can you guarantee the set $AA + AA$ to be? This is actually the context in which this problem was initially posed. Questions like these naturally arise in the study of additive number theory, see [**34**]. The subject matter contained in this chapter then developed after it became apparent that it could be analyzed like the Erdős distance problem. *Hint:* Consider the Cartesian product setting.

Chapter 10

Vector spaces over finite fields

Now we introduce the very basics of finite fields, to illustrate another way that the basic ideas that have already been presented can be extended to study other types of problems. The structure of fields used in this book is only the tip of the iceberg. Starting with a formal definition of a field would be quite cumbersome, so it is probably more natural to think of a field as a system that works like numbers with identities, division, and commutative addition and multiplication. Some examples of fields are the real numbers, and the complex numbers. In this chapter and the next, i denotes the square root of -1.

Exercise 10.1. Just from the cursory definition of a field given in the above paragraph, why are the integers not a field? *Hint*: The set of nonnegative integers are not a field because they do not have additive inverses. Even if we consider all the integers, what is still missing?

1. Finite fields

In this book, we are focusing on finite fields. The *order* of a finite field is the number of elements in it. So the real and complex numbers could not be finite fields, as they have infinite size. An example of a finite field would be what we call $\mathbb{Z}_5$, or the numbers 0 through 4, where you treat the numbers like a clock. That is to say, if you add 3 and 4, you get 2. That's because $3 + 4 = 7$, and $7 - 5 = 2$. The algorithm is as follows: add or multiply as usual, and if you get a number not between 0 and 4, add or subtract multiples of 5 until you are between 0 and 4. This phenomenon is often written as $3 + 4 \equiv 2$, or 3+4 is *congruent* to 2 *modulo* or *mod* 5.

One curious thing about $\mathbb{Z}_5$ is that, in some sense, -1 is a square. To see this, note that 4 behaves like -1, in that it is the element that represents $0 - 1$. Then we recall that 4 is a perfect square, namely 2^2. So we can think of 2 as $\sqrt{-1}$.

Exercise 10.2. Show by hand that $\mathbb{Z}_7$ has no $\sqrt{-1}$.

We call two special elements of our field *identities*. There is the *multiplicative identity*, which is usually denoted 1, just as in the more commonplace fields. This is because anything times 1 is itself again. Then the *additive identity* is usually written as 0, for similar reasons. We also guaranteed *inverses*. The *multiplicative inverse* of an element, a, is the element, b, such that $a \cdot b = 1$ in the field. *Additive inverses* are defined similarly. Note that 0 cannot have a multiplicative inverse. If we want to discuss the non-zero elements of a given finite field, that is, the elements with multiplicative inverses, we often denote it $\mathbb{F}_q^*$.

A curious thing that might pique your attention, is the restriction to finite fields involving prime numbers. It might seem odd, that in a highly geometric book, primality could matter at all, however, in order for some of the fundamental properties of fields to hold, we need their orders to be powers of primes. In order to illustrate an example of why fields must have this restriction, try the following exercises. In this Chapter, we simplify things by dealing with fields of prime order, or where the power of the prime is just one.

Exercise 10.3. Show that every number has a multiplicative inverse in $\mathbb{F}_{11}$. For example, $3 \cdot 4 = 12 = 1 + 11$. This means $3 \cdot 4 \equiv 1$. So 3 has the multiplicative inverse 4, and 4 has the multiplicative inverse 3.

Exercise 10.4. Show that some numbers do not have a multiplicative inverse in $\mathbb{F}_{12}$.

There are plenty of different ways to think of finite fields, and an abundance of rich theory that goes deep and far in many different directions. However, for the purposes of this book, the basic ideas presented here are probably enough. For completeness' sake only, we include a formal definition of a field:

Definition 10.1. A *field* $\mathbb{F} = (F, +, \cdot)$ is a set, F, with two unique special elements, 1 and 0, and two functions that satisfy the following conditions:

$+ : F \times F \to F$.
$\cdot : F \times F \to F$.
(i) $+(x, y) = +(y, x)$ for all $x, y \in F$.
(ii) $\cdot(x, y) = \cdot(y, x)$ for all $x, y \in F$.
(iii) $+(x, 0) = x$ for all $x \in F$.
(iv) $\cdot(x, 1) = x$ for all $x \in F\backslash\{0\}$.
(v) $\cdot(x, 0) = 0$ for all $x \in F$.
(vi) For all $x \in F$, there exists a unique $y \in F$ such that $+(x, y) = 0$.
(vii) For all $x \in F\backslash\{0\}$, there exists a unique $y \in F$ such that $\cdot(x, y) = 1$.
(viii) $+(x, +(y, z)) = +(+(x, y), z)$ for all $x, y, z \in F$.
(ix) $\cdot(x, \cdot(y, z)) = \cdot(\cdot(x, y), z)$ for all $x, y, z \in F$.
(x) $\cdot(x, +(y, z)) = +(\cdot(x, y), \cdot(x, z))$, for all $x, y, z \in F$.

2. Vector spaces

Now that you have an idea of what a finite field is, we will turn our attention to *vector spaces*. This term is basically just a fancy way of indicating that we are using the ideas behind the Cartesian coordinate system, but on something other than real numbers. In general, vector spaces can be applied over many things. In this book, we are dealing with vector spaces over finite fields. That means that we have a Cartesian coordinate system, but with on a finite set. Figure 10.1 is a popular representation.

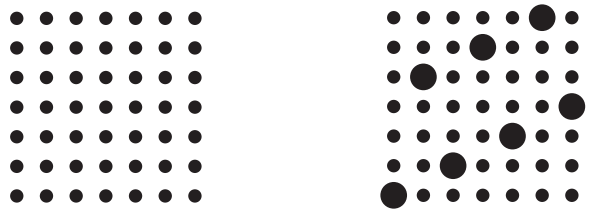


Figure 10.1. On the left is one way of visualizing $\mathbb{F}_7^2$. On the right, the larger points represent the line through the origin generated by the element $(2, 1)$.

On the surface, this looks just like any other grid. However, this grid has the property that if you walk off of the top, you end up on the bottom. The same is true from left to right. At first it might just seem to behave like any number of popular video games from the eighties, but this little detail ends up providing plenty of arithmetic pitfalls of its own!

How could you model "walking" through this grid? What would a line look like here? In the plane, one way to specify a line is with a point and a slope. So we will try to find the most sensible way to specify a line in the finite fields setting by a point and a slope. Define the line l_x, in $\mathbb{F}_q^2$, which passes through the origin, as follows:

$$l_x := \{p \in \mathbb{F}_q^2 : p = tx, \text{where } t \in \mathbb{F}_q\}.$$

Exercise 10.5. In $\mathbb{F}_7^2$, which points belong to the line through the element $(2, 2)$ with slope $(1, 3)$?

Obviously, we have introduced this topic because it should have something to do with the Erdős distance problem. Now, since we cannot be sure that the square root is defined for every element, the standard Euclidean distance will not work. If we think about what kinds of features of metrics would be most necessary for the finite field setting, we might recall the concept of homogeneity, introduced in Chapter 1. We said that if we measured a stick in one location, then took it somewhere else and measured it again, we would want to get the same measurement. In vector spaces over finite fields, this means that two points that form a given configuration with respect to one another, then if we move that configuration somewhere else in the space, or rotate it somehow, that the "distance" between the two points remain the same. Another way of saying this is that we will want our notion of distance to behave well under *rigid motions*, or rotations and translations.

Figure 10.2 will illustrates how the first pair of points is translated and rotated to form the second and third point pairs. The first becomes the second by translating to the right by 14 elements, and up by 6 elements. The first becomes the third by translating to the right by 5 elements and up by twelve, and rotating 60° clockwise.

We want to make very clear that the object we will introduce and loosely call "distance" is not actually a metric in the strict sense, as it does not obey the triangle inequality. This is not really an issue though, as a finite field is not ordered[1], so any notion of metric we could invent would immediately violate the triangle inequality.

Now, if the theory of finite fields is new to you, you might be wondering why finite fields are not ordered. This is because they "wrap around". So any intuitive notion of greater than or less than would break down when you add one to the largest element. Again, this is all restricted to the case were the finite field is of prime order. Things would get really messy if we tried to impose an ordering on other finite fields.

The generally accepted notion of distance in a vector space over a finite field looks quite similar to the Euclidean metric elsewhere, without the square root. If x and y are two points in $\mathbb{F}_q^2$, we define their distance as follows:

$$\|x - y\| = (x_1 - y_1)^2 + (x_2 - y_2)^2.$$

[1] The order of a finite field refers to its size. The concept of *ordering* is completely different.

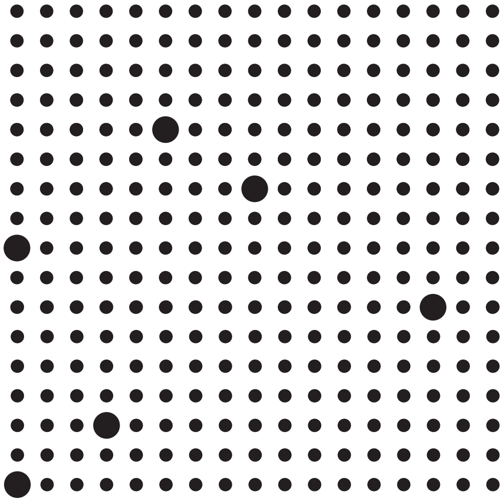


Figure 10.2. Here we consider three pairs of points in $\mathbb{F}_{17}^2$. The lower left pair is $(0,0)$ and $(3,2)$. The second pair, $(14,6)$ and $(0,8)$, appears to be split. The upper pair is $(5,12)$ and $(8,10)$.

Of course, this generalizes to d dimensions, in $\mathbb{F}_q^d$ as:

$$\|x-y\| = \sum_{j=0}^{d}(x_j - y_j)^2.$$

Now, we have made quite a fuss about this object behaving well under translations and rotations. Translations are relatively easy to imagine, but rotations are a bit more complicated to describe. Rather than go into all of the details of what kinds of objects are analogous to rotations would take a lot of time, and take us too far off course. We would have to deal with the special orthogonal group on a vector space over a finite field. So, in Proposition 10.1, we will just show that this distance is invariant under translations, and if you are really interested, you can go through Exercise 10.7 to see an example of a rotation in a vector space over a finite field, and verify that the distance is preserved.

Proposition 10.1. *The generally accepted notion of distance in a vector space over a finite field is invariant under translations.*

Proof. Given two points x and y in $\mathbb{F}_q^d$, and a translation, T, also in $\mathbb{F}_q^d$, we need to show that $\|x-y\| = \|x'-y'\|$, where $x' = x+T$, and $y' = y+T$. Let x have the coordinates $(x_1, x_2, ..., x_d)$, and denote the coordinates for x', y, y', and T similarly.

$$\begin{aligned}\|x-y\| &= \sum_{j=0}^{d}(x_j - y_j)^2 \\ &= \sum_{j=0}^{d}(x_j + T_j - T_j - y_j)^2 \\ &= \sum_{j=0}^{d}((x_j + T_j) - (y_j + T_j))^2 \\ &= \sum_{j=0}^{d}(x'_j - y'_j)^2 \\ &= \|x' - y'\|\end{aligned}$$

□

Exercise 10.6. Recall the point pairs in $\mathbb{F}_{17}^2$ depicted in Figure 10.2. The lower left pair is $(0,0)$ and $(3,2)$, the second pair is $(14,6)$ and $(0,8)$, and the upper pair is $(5,12)$ and $(8,10)$. Show that each for each pair of points, the two points have a distance of 13 from one another. *Hint*: Use the proof of Proposition 10.1.

Now we will introduce the notion of a sphere or circle for vector spaces over finite fields. In the vector space over the reals, $\mathbb{R}^d$, we define a circle as all of the points that are a particular distance from a given point. We will do the same here. Let S_j denote the sphere of radius $j \in \mathbb{F}_q$ centered at the origin in $\mathbb{F}_q^d$ as:

$$S_j = \{x \in \mathbb{F}_q^d : \{\|x\| = j\}.$$

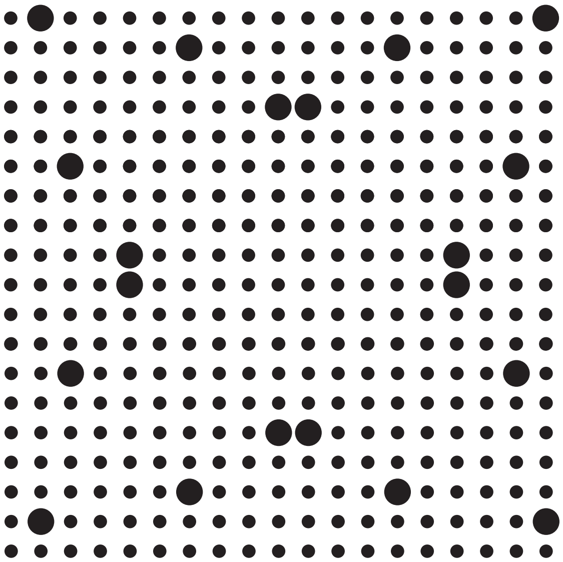


Figure 10.3. This is one way of visualizing the circle of radius 2, centered at the origin, in $\mathbb{F}_{19}^2$.

You can similarly define the sphere of radius j centered at a point y by $\{x \in \mathbb{F}_q^d : \{\|x - y\| = j\}$. Since this definition is quite abstract, we have Figure 10.3 to show what a circle of radius 2, centered at the origin looks like in $\mathbb{F}_{19}^2$.

We know that rotations around a point, p, take points on a circle of a given radius, centered at p to points on the same circle. Now, as promised, is an example of a rotation in a vector space over a finite field.

Exercise 10.7. Consider $\mathbb{F}_{17}^2$, and the 2×2 matrix:

$$R = \begin{pmatrix} 3 & -3 \\ 3 & 3 \end{pmatrix}$$

.

Now, consider the point $a = (0, 2)$. Treat this point as a non-square matrix, and use matrix multiplication to check that:

$$Ra = \begin{pmatrix} 3 & -3 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \end{pmatrix} = (-6, 6)$$

.

Now check that the distance to the origin is unchanged. By that we mean show that:

$$\|(0, 2)\| = \|(-6, 6)\|.$$

This means that both a and Ra are on the circle of radius 4 in $\mathbb{F}_{17}^2$, so R makes sense as a rotation.

3. Exponential sums in finite fields

The Fourier transform is an important part of any mathematician's toolkit. It is very powerful, and can be used in many different ways. Here, we confine ourselves to the finite field setting. Although it is often introduced on the real numbers first, we believe that the fundamental ideas behind it make just as much sense here, if not more.

Before we get to a formal definition of the Fourier transform, we will gently introduce some surrounding ideas, to make the transition of reasoning easier. First, consider what happens if you sum up all of the elements in a finite field.

$$\sum_{x \in \mathbb{F}_q} x = 0.$$

Why is that? Well, since every non-zero element has one unique additive inverse, each such element and its inverse sum to zero. Then of course, the only remaining element in the sum is zero itself.

Recall that the kth roots of unity can be written as:

$$e^{\frac{2\pi i \cdot 0}{k}} = 1, e^{\frac{2\pi i}{k}}, e^{\frac{2\pi i \cdot 2}{k}}, ..., e^{\frac{2\pi i (k-1)}{k}}.$$

What happens when we sum all of the roots of unity of a given order?

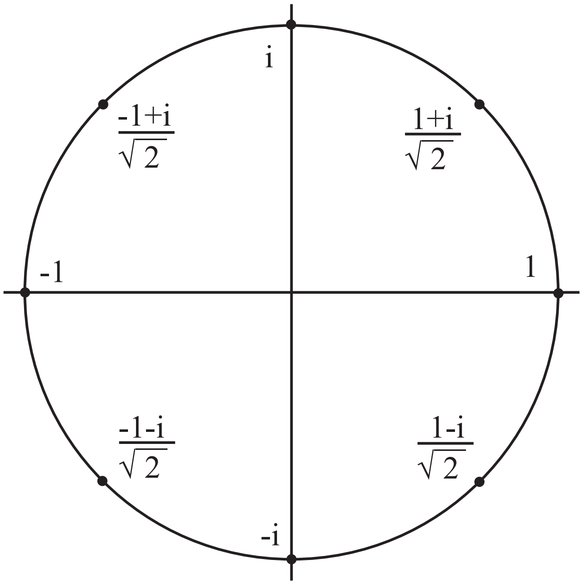


Figure 10.4. The points represent the eighgth roots of unity in the complex plane.

$$\sum_{j=0}^{k} e^{\frac{2\pi i j}{k}} = 0.$$

This can be seen by vector addition. If you are not convinced, do the following exercise.

Exercise 10.8. Show that the sum of the fifth roots of unity is 0 two different ways.
i) Do this by writing each root of unity as a real number plus an imaginary number, e.g. $x + yi$. Then sum them up in that form to verify that the real and imaginary parts of the sum are both zero.
ii)Now draw the vectors for each of the roots of unity head to tail, in any order, and notice that they form a closed path. Therefore their vector sum is zero.

Note that the following is also true:

$$\sum_{j=0}^{k-1} e^{\frac{2\pi i j a}{k}} = 0. \tag{10.2}$$

This is because we can take a kth root of unity, $e^{\frac{2\pi i a}{k}}$, and rotate it $\frac{ja}{k}$ of the way around the unit circle by multiplying it by itself j times. The fact that this sums to zero is an example of a property called *orthogonality.* You can see this by doing the next exercise. Now, in this section, we will only show orthogonality in one dimension, but soon enough, we will employ orthogonality in more dimensions. It will follow for the same logical reasons.

Exercise 10.9. Show explicitly by hand that (10.2) holds for $k = 7$ and $a = 2$. Notice what happens to each term.

Exercise 10.10. Show explicitly by hand that the following sum over two dimensions, represented by j and j', is zero for nonzero elements a, and is q^2 if $a = 0$.

$$\sum_{j,j'=0}^{k-1} e^{\frac{2\pi i (j+j')a}{k}}$$

Hint: Try seperating the sum as follows:

$$\sum_{j=0}^{k-1}\sum_{j'=0}^{k-1} e^{\frac{2\pi i j a}{k}} e^{\frac{2\pi i j' a}{k}} = \sum_{j=0}^{k-1}\left(e^{\frac{2\pi i j a}{k}} \left(\sum_{j'=0}^{k-1} e^{\frac{2\pi i j' a}{k}} \right)\right).$$

Then use orthogonality in each sum seperately.

That was not so bad. Now, we will turn up the heat a little bit and consider a different sort of sum. We first need to consider a special kind of function called an *additive character.* This function takes elements in whichever field we are considering, and maps them into roots of unity in the unit circle in the complex plane. We can formally define an additive character, χ in the following way:

$$\chi : \mathbb{F}_q \to \mathbb{C}.$$

$$\chi(a) = e^{\frac{2\pi i a}{q}}, a \in \mathbb{F}_q$$

Notice what happens if $a = 0$, $\chi(0) = e^{\frac{2\pi i \cdot 0}{q}} = 1$.

So we can rewrite (10.2) in terms of our additive character, χ, for any nonzero element a.

$$\sum_{j=0}^{k-1} \chi(ja) = \sum_{j=0}^{k-1} e^{\frac{2\pi i j a}{k}} = 0. \tag{10.3}$$

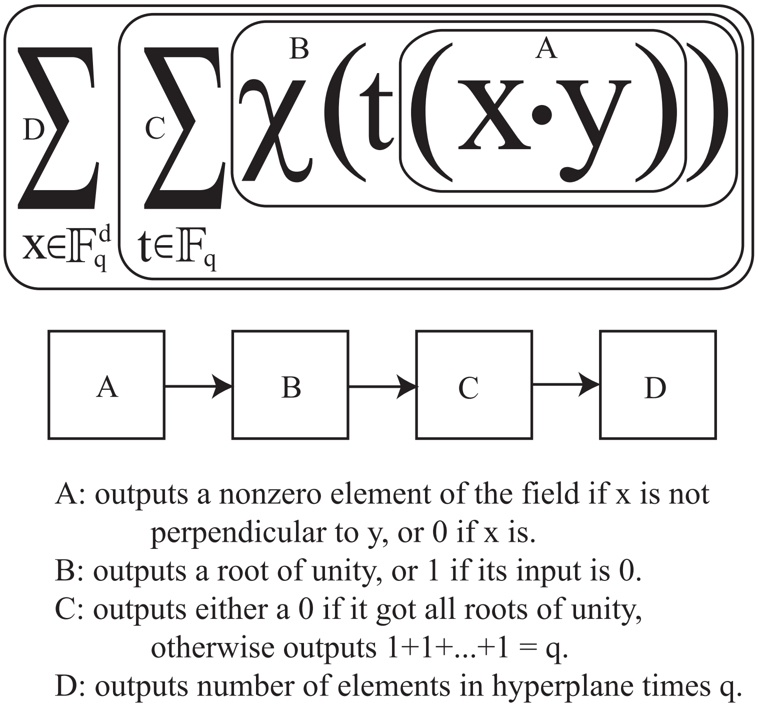
However, if $a = 0$, $\chi(ja) = 1$ for every j. So in that case, (10.3) will look like this:

$$\sum_{j=0}^{k-1} \chi(j0) = \sum_{j=0}^{k-1} e^{\frac{2\pi i j(0)}{k}} = \sum_{j=0}^{k-1} 1 = q. \tag{10.4}$$

This turns out to be extraordinarily useful when we try to count things. Since we have just defined lines and circles, we will use this new device to count how many points are on a line, or in a circle. Of course, those objects exist in vector spaces over finite fields, so we will have to find a way to make our additive character make sense in higher dimensions. Since our additive character only takes elements of the one-dimensional finite field as input, we will need to find a way to bring elements in the vector space to our finite field. Keep that in mind as you carefully examine the following expression. Below, y is some fixed vector in $\mathbb{F}_q^d$, and $\cdot$ denotes the usual dot product.

$$q^{-1} \sum_{x \in \mathbb{F}_q^d} \sum_{t \in \mathbb{F}_q} \chi(t(x \cdot y)) \tag{10.5}$$

Now, at first blush, this might not scream about counting the number of points in a line, but hopefully it will eventually. Of course, it is more complicated than (10.4), but that is the heart of it. We will offer two different explanations of this expression. Please read through both of them very carefully. Sometimes all it takes is another viewpoint, and everything becomes clear. Also, for the next little while, do not worry about computing these sums, just worry about interpreting their meanings.


D
C
B
A
$\sum_{x\in\mathbb{F}_q^d} \sum_{t\in\mathbb{F}_q} \chi(t(x\cdot y))$
A
B
C
D

A: outputs a nonzero element of the field if x is not perpendicular to y, or 0 if x is.

B: outputs a root of unity, or 1 if its input is 0.

C: outputs either a 0 if it got all roots of unity, otherwise outputs 1+1+...+1 = q.

D: outputs number of elements in hyperplane times q.

Figure 10.5. The heart of (10.5). Each part of the sum can be viewed as a box in the machine. Each box takes starts with input and gives output to the next box. The end output must be scaled by q^{-1} to be accurate, but the main ideas are here.

As you can see, when $(x \cdot y)$ is nonzero, the sum over t is 0, so the whole inner sum becomes zero. So we only need to consider the terms where the $(x \cdot y)$ is zero. If $(x \cdot y) = 0$ though, then the sum over t is q, as in (10.4). This means that for each point that is perpendicular to the vector y, the sum over t returns q. This explains the factor of q^{-1} out front. It scales the sum so that each element in the hyperplane returns a 1 and not a q. So we get that y is a normal vector to a $d-1$ dimensional hyperplane, and our sum counts the number of points in it.

If the first explanation was not enough, the following explanation is more like an assembly line, or a computer program. Figure 10.5 shows how each of the main components of our "counting machine" fit together. We view part A as testing each x against the given y. Part B yields a different result depending on what part A spat out. Then part C sums the outputs of part B to yield zero if x is not perpendicular to y, and q if they are. Part D does this for every $x \in \mathbb{F}_q^d$. Of course you still have to scale when all is said and done.

To ensure that this makes sense, we will show you how to count something else. As you may have guessed, we will also have to make sense of circles and spheres. We will show you how to count the number of elements in a circle, but you will count the number of elements in a $d-1$ dimensional sphere in Exercise 10.11.

Recall the ideas behind (10.5). If we forget about the fact that we were counting elements in a hyperplane, and just think about how we counted elements in some special set, the reasoning would go as follows. We summed over the vector space to ask each element whether or not it belonged in our set. Then for each element in this big sum, we ran our additive character sum, which returned a 0 if the element was not in our set, and a q if it was, then we scaled everything by q^{-1} return 1 for each element in our set.

The next expression will use similar ideas to count the number of elements on a circle of radius r in $\mathbb{F}_q^2$.

$$q^{-1} \sum_{x \in \mathbb{F}_q^2} \sum_{t \in \mathbb{F}_q} \chi(t(x_1^2 + x_2^2 - r)) \tag{10.6}$$

Notice that the basic setup for (10.5) is the same as it was for (10.6), but in the additive character, we have a different expression. Think about which elements in $\mathbb{F}_q^2$ will yield an argument or input of 0 for χ.

$$\begin{aligned} x_1^2 + x_2^2 - r &= 0 \\ x_1^2 + x_2^2 &= r \\ \|x\| &= r \end{aligned}$$

So χ will get a zero input only when x is on the circle of radius r centered at the origin. How would you modify (10.6) to count the number of elements on a circle of radius r centered at a point $y \in \mathbb{F}_q^d$?

$$q^{-1} \sum_{x \in \mathbb{F}_q^2} \sum_{t \in \mathbb{F}_q} \chi(t((x_1 - y_1)^2 + (x_2 - y_2)^2 - r)) \tag{10.7}$$

As with (10.6), χ only gets in zeros when $\|x - y\| = r$, or when x is on a circle of radius r from y.

Exercise 10.11. Use the tools that you have learned with the hyperplane counting sum in d dimensions, (10.5), and the circle counting sums in 2 dimensions, (10.6), and (10.7), to construct a sum that counts the number of elements on a $d - 1$ dimensional sphere in d dimensions, centered at some $y \in \mathbb{F}_q^d$. Again, note that we do not expect you to compute these sums yet.

We just have one last thing before we move on to the next section, and it is simpler than the previous things. Think of it as a cooldown. If we specify a subset $E \subset \mathbb{F}_q^d$, then we can count the number of elements in the subset by using a special function called the *indicator function* or *characteristic function* of E. We will denote this function $E(x)$. It takes the value 1 if $x \in E$ and 0 if $x \notin E$. Consider the following sum:

$$\sum_{x \in \mathbb{F}_q^d} E(x)$$

.

It will run through every element in $\mathbb{F}_q^d$ and add a 1 if the element is in the set, and add nothing if not. So, we can be assured that:

$$\sum_{x \in \mathbb{F}_q^d} E(x) = \#E,$$

by definition. Now that you are aquainted with characteristic functions and counting the number of elements in a particular object, we will make things just a bit more complex by throwing in "weights" for the elements. Although this is a mildly misleading analogy as it stands, the following section will reveal its purpose. We will consider (10.6) again, but this time, instead of counting each point once, we will count different points different numbers of times. Suppose we wanted to know how many points of the circle are in a particular subset $E \subset \mathbb{F}_q^d$? Well, we could multiply each term in the sum by the characteristic function of E, and then only add one when the element under consideration is both in E and the circle. The end result would look like this:

$$(10.8) \qquad q^{-1} \sum_{x \in \mathbb{F}_q^2} \left(E(x) \sum_{t \in \mathbb{F}_q} \chi(t(x_1^2 + x_2^2 - r)) \right)$$

So far we have only weighted our terms by 1 or 0. The next section will weight each term by an arbitraty function defined on $\mathbb{F}_q^d$.

4. The Fourier transform

Now that you have the basic idea of counting things with exponential sums, we will introduce the *Fourier transform.* It is one of the most important and fundamental tools in mathematics. If you plan to do mathematics, chances are that you will end up using this quite often.

Definition 10.9. Let f be a function on $\mathbb{F}_q^d$. For $m \in \mathbb{F}_q^d$, let

$$\widehat{f}(m) = q^{-d} \sum_{x \in \mathbb{F}_q^d} e^{-\frac{2\pi i}{q} x \cdot m} f(x) = q^{-d} \sum_{x \in \mathbb{F}_q^d} \chi(-x \cdot m) f(x).$$

The minus sign in the definition is there for reasons that we will not get into in this book, but the rest should appear somewhat familiar. Now in the Fourier transform, we do not have the luxury of seperating the sum into an element picking sum and an element testing sum, or parts C and D of our machine in Figure 10.5, respectively. This makes it difficult to get as clean of a "physical" interpretation of this particular device. However, hopefully it does not appear too intimidating after dealing with the simpler exponential sums.

Now, to further aquaint you with the Fourier transform, we will guide you through a basic calculation. This will give you a feel for the kinds of computations that lay ahead. The first is called *Fourier inversion.* If you know the Fourier transform of a function everywhere, you can construct the original function using this method.

$$f(x) = \sum_{m \in \mathbb{F}_d^q} e^{\frac{2\pi i}{q} x \cdot m} \widehat{f}(m). \tag{10.10}$$

To see this, start with the definition of the Fourier transform and work backwards.

$$\sum_{m \in \mathbb{F}_d^q} e^{\frac{2\pi i}{q} x \cdot m} \widehat{f}(m) = \sum_{m \in \mathbb{F}_d^q} e^{\frac{2\pi i}{q} x \cdot m} \left(q^{-d} \sum_{y \in \mathbb{F}_q^d} e^{-\frac{2\pi i}{q} y \cdot m} f(y) \right),$$

by definition of the Fourier transform of f at each y.

$$\sum_{m \in \mathbb{F}_d^q} e^{\frac{2\pi i}{q} x \cdot m} \left(q^{-d} \sum_{y \in \mathbb{F}_q^d} e^{-\frac{2\pi i}{q} y \cdot m} f(y) \right) = q^{-d} \sum_{m \in \mathbb{F}_d^q} e^{\frac{2\pi i}{q} x \cdot m} \left(\sum_{y \in \mathbb{F}_q^d} e^{-\frac{2\pi i}{q} y \cdot m} f(y) \right),$$

by factoring out q^{-d}.

$$q^{-d} \sum_{m \in \mathbb{F}_d^q} e^{\frac{2\pi i}{q} x \cdot m} \left(\sum_{y \in \mathbb{F}_q^d} e^{-\frac{2\pi i}{q} y \cdot m} f(y) \right) = q^{-d} \sum_{m \in \mathbb{F}_d^q} \left(\sum_{y \in \mathbb{F}_q^d} e^{\frac{2\pi i}{q} x \cdot m} e^{-\frac{2\pi i}{q} y \cdot m} f(y) \right),$$

by moving the exponential into the sum.

$$q^{-d}\sum_{m\in\mathbb{F}_d^q}\left(\sum_{y\in\mathbb{F}_q^d} e^{\frac{2\pi i}{q}x\cdot m}e^{-\frac{2\pi i}{q}y\cdot m}f(y)\right)=q^{-d}\sum_{m\in\mathbb{F}_d^q}\left(\sum_{y\in\mathbb{F}_q^d} e^{\frac{2\pi i}{q}(x-y)\cdot m}f(y)\right),$$

by adding the exponents.

$$q^{-d}\sum_{m\in\mathbb{F}_d^q}\left(\sum_{y\in\mathbb{F}_q^d} e^{\frac{2\pi i}{q}(x-y)\cdot m}f(y)\right)=q^{-d}\sum_{y\in\mathbb{F}_d^q}\left(\sum_{m\in\mathbb{F}_q^d} e^{\frac{2\pi i}{q}(x-y)\cdot m}f(y)\right),$$

by switching the order of summation, which is fine as everything is finite here. Due to orthogonality, this sum is only nonzero when $x = y$, at which point it returns the value of $f(y)$, which is of course $f(x)$, q^d times. So,

$$q^{-d}\sum_{y\in\mathbb{F}_d^q}\left(\sum_{m\in\mathbb{F}_q^d} e^{\frac{2\pi i}{q}(x-y)\cdot m}f(y)\right)=q^{-d}(q^d f(x))=f(x),$$

as promised. Now, we chose to write this calculation out sowly and step-by-step so you could soak up every bit of reasoning employed. Please make sure that no steps are mysterious, as all of these ideas will be taken for granted in the next chapter.

Before we continue, we will remind you of a few basic concepts in complex space. As is usually the case, we will let $\overline{z}$ denote the *complex conjugate* of z. So if

$$z = x + yi = re^{i\theta},$$

then its complex conjugate will be written

$$\overline{z} = x - yi = re^{-i\theta}.$$

When a modulus is taken in complex space, we can think of it as $|z|^2 = z\overline{z}$. Now we are ready to introduce the Plancherel formula.

$$\sum_{m \in \mathbb{F}_q^d} \left|\widehat{f}(m)\right|^2 = q^{-d} \sum_{x \in \mathbb{F}_q^d} |f(x)|^2 \tag{10.11}$$

The proof of the Plancherel formula is also an elementary calculation. We will leave it as an exercise.

Exercise 10.12. Prove the Plancherel formula, (10.11), by writing the modulus as a product of $f(x)\overline{f(x)}$, and use the proof of the Fourier inversion formula.

Chapter 11

Distances in vector spaces over finite fields

In this chapter, we are going to study the Erdős distance problem in vector spaces over finite fields. Even though we defined everything necessary in the last chapter, we will repeat some definitions just to make them stick a little better, and to reduce the initial amount of page flipping, so that you can keep track of what is really going on.

1. The setup

Let $\mathbb{F}_q$ denote a finite field with q elements. For the sake of clarity, we shall confine our attention to the case where q is a prime number. We also assume, for the sake of computational simplicity, that -1 is not a square in $\mathbb{F}_q$ in the sense that there does not exist $s \in \mathbb{F}_q$ such that $s^2 = -1$. Let $\mathbb{F}_q^d$, $d \geq 2$, denote the d-dimensional vector space over $\mathbb{F}_q$. What form does the Erdős distance problem take in this setting? Given $E \subset \mathbb{F}_q^d$, let

$$\Delta(E) = \{\|x - y\| : x, y \in E\},$$

where we define distance as before,

$$\|x - y\| = (x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_d - y_d)^2.$$

It is tempting to conjecture, as before, that

$$\#\Delta(E) \gtrapprox (\#E)^{\frac{2}{d}}.$$

Unfortunately, this is just not true! Observe that if $E = \mathbb{F}_q^d$, $\#E = q^d$, whereas $\#\Delta(E) = q$. It follows that, in general, the best estimate we can expect is

$$\#\Delta(E) \gtrsim (\#E)^{\frac{1}{d}}.$$

At least in two dimensions, this estimate is fairly easy to achieve (see Exercise 6.1 below), so is this the end of the story? Fortunately, the answer is no. In [**4**], the following result is proved.

Theorem 11.1. *Let $E \subset \mathbb{F}_q^2$ such that $\#E = q^{2-\epsilon}$. Then there exists $\delta = \delta(\epsilon)$ such that*

$$\#\Delta(E) \gtrsim (\#E)^{\frac{1}{2}+\delta}. \tag{11.1}$$

The proof of this result is beyond the scope of this book. The goal of this Chapter is to prove a non-trivial version of (11.1) and to clarify the nature of the exponents. We shall prove the following result, which is from [**18**].

Theorem 11.2. *Let $E \subset \mathbb{F}_q^d$, $d \geq 2$, such that $\#E \gtrsim q^{\frac{d+1}{2}}$. Then*

$$\#\Delta(E) \gtrsim q. \tag{11.2}$$

The exponent $\frac{d+1}{2}$ is sharp in the following sense: for every $\epsilon > 0$, there exists a set, E, of size approximately $q^{\frac{d+1}{2}-\epsilon}$ for which the size of the distance set E is $\lesssim q^{1-\delta}$, where δ is a function ϵ. This argument is presented in [**19**].

To prove Theorem 11.2, consider

$$\#\{(x,y) \in E \times E : \|x-y\| = j\}$$

for some $j \in \mathbb{F}_q$, $j \neq 0$. Let $E(x)$ denote the characteristic function of E, the function which equals 1 if $x \in E$ and 0 otherwise. Let $S_j(x)$ denote the characteristic function of the sphere $\{x \in \mathbb{F}_q^d : \{\|x\| = j\}$. Remember, that since the "distance" is defined differently, a sphere in a vector space over a finite field will probably not superficially resemble a sphere in Euclidean space. We have

$$\#\{(x,y) \in E \times E : |x-y|^2 = j\} = \sum_{x,y \in \mathbb{F}_q^d} E(x)E(y)S_j(x-y). \tag{11.3}$$

In order to proceed, we will remind you of the definition of the Fourier transform in this setting.

If f is a function on $\mathbb{F}_q^d$, and $m \in \mathbb{F}_q^d$, let

$$\widehat{f}(m) = q^{-d} \sum_{x \in \mathbb{F}_q^d} e^{-\frac{2\pi i}{q} x \cdot m} f(x).$$

We will need to recall a few basic facts from the last chapter. First, recall the technique of Fourier inversion.

$$f(x) = \sum_{m \in \mathbb{F}_d^q} e^{\frac{2\pi i}{q} x \cdot m} \widehat{f}(m).$$

Next, recall the Plancherel formula:

$$\sum_{m \in \mathbb{F}_q^d} |\widehat{f}(m)|^2 = q^{-d} \sum_{x \in \mathbb{F}_q^d} |f(x)|^2.$$

It follows that the right hand side of (11.3) equals

$$\sum_{x,y,m \in \mathbb{F}_q^d} E(x)E(y)e^{\frac{2\pi i}{q}(x-y)\cdot m} \widehat{S}_j(m) = q^{2d} \sum_{m \in \mathbb{F}_q^d} |\widehat{E}(m)|^2 \widehat{S}_j(m). \tag{11.4}$$

Now you have most of the tools necessary to explore the Erdős distance problem in vector spaces over finite fields. Good luck!

2. The argument

This section is the last of the main part of the book, and as such, is much denser in content, and therefore more likely to be difficult. We hope that you have enjoyed the book thus far, and see this section as a kind of parting gift. If it does not all sink in immediately, do not worry. This section is intended to give you something to work on for a long time to come.

Lemma 11.3. *With the notation above, if $j \neq 0$, then*

$$|\widehat{S}_j(m)| \lesssim q^{-\frac{d+1}{2}},$$

and

$$\#S_j \approx q^{d-1}.$$

Assume Lemma 11.3 for a moment. The right hand side of (11.4) equals

$$q^{2d}|\widehat{E}(0,\ldots,0)|^2 \widehat{S}_j(0,\ldots,0) + q^{2d} \sum_{m \neq (0,\ldots,0)} |\widehat{E}(m)|^2 \widehat{S}_j(m) = I + II.$$

The first term is the same sum as above, but in the special case that $m = (0,0,...,0)$. Henceforth, we call it I. The sum over $m \neq (0,0,...,0)$ is called II.

Now,

$$I = q^{2d}q^{-2d}(\#E)^2 q^{-d}\#S_j \approx (\#E)^2 q^{-1}.$$

Because I is a positive real number, $|II|$ will be less than the right hand side of (11.4). Now appeal to Lemma 11.3 and the Plancherel formula to see that

$$\begin{aligned} |II| &\lesssim q^{2d} \sum_{m \in \mathbb{F}_q^d} \left|\widehat{E}(m)\right|^2 \left|\widehat{S_j}(m)\right| \\ &\lesssim q^{2d} q^{-\frac{d+1}{2}} \sum_{m \in \mathbb{F}_q^d} |\widehat{E}(m)|^2 \\ &= q^{\frac{d-1}{2}} \#E. \end{aligned}$$

Since

$$\sum_j \#\{(x,y) \in E \times E : \|x-y\| = j\} = (\#E)^2,$$

it follows that

$$\#\Delta(E) \gtrsim \min\left\{q, \frac{\#E}{q^{\frac{d-1}{2}}}\right\},$$

as desired.

In order to prove Lemma 11.3 we need the following preliminary result about Gauss sums.

Lemma 11.4. *Let* $G(m,k)=\sum_{x\in\mathbb{F}_q^d} e^{\frac{2\pi i(x\cdot m-k|x|^2)}{q}}$. *Then if* $k\neq 0$,

$$G(m,k)=e^{\frac{2\pi i|m|^2}{4kq}}g^d(k), \tag{11.5}$$

$$g(k)=\pm i\sqrt{q}, \tag{11.6}$$

and, consequently,

$$g^d(k)=(\pm i)^d\cdot q^{\frac{d}{2}}, \tag{11.7}$$

where $g(k)$ *is the "standard" Gauss sum*

$$g(k)=\sum_{x_j\in\mathbb{F}_q} e^{\frac{2\pi ikx_j^2}{q}}.$$

To prove Lemma 11.4, we write

$$\sum_{x_j\in\mathbb{F}_q} e^{\frac{2\pi i(m_jx_j-kx_j^2)}{q}}=e^{\frac{2\pi im_j^2}{4kq}}\sum_{x_j\in\mathbb{F}_q} e^{-\frac{2\pi ik(x_j-m_j/2k)^2}{q}},$$

just by comlpleting the square. Be sure to check that this works! After that, to see the next step, just notice that summing over all elements in a field, or all elements in a field in a different order is the same. This is a kind of change of variables. This is illustrated in Exercise 11.1, right after the proof. This means that

$$e^{\frac{2\pi im_j^2}{4kq}}\sum_{x_j\in\mathbb{F}_q} e^{-\frac{2\pi ik(x_j-m_j/2k)^2}{q}}=e^{\frac{2\pi im_j^2}{4kq}}g(k),$$

and the identity (11.5) follows.

Exercise 11.1. Show the following equality holds when m and k are some elements in $\mathbb{F}_q$.

$$\sum_{x\in\mathbb{F}_q} e^{-\frac{2\pi ik(x-m/2k)^2}{q}}=\sum_{y\in\mathbb{F}_q} e^{-\frac{2\pi iky^2}{q}}$$

Think of it as a change of variables. Since the sum is still taken over all elements, it is the same sum!

We now prove (11.6) and (11.7). Indeed,

$$|g(k)|^2 = \sum_{u,v \in \mathbb{F}_q} e^{\frac{2\pi i k(u^2-v^2)}{q}}$$

$$= \sum_{t \in \mathbb{F}_q} e^{\frac{2\pi i k t}{q}} n(t),$$

where

$$n(t) = \#\{(u,v) \in \mathbb{F}_q \times \mathbb{F}_q : u^2 - v^2 = t\}.$$

Lemma 11.5. *We have $n(0) = 2q-1$, and $n(t) = q-1$ if $t \neq 0$.*

Write $u^2 - v^2 = (u-v)(u+v)$. Since $u-v$ and $u+v$ determine u and v uniquely, it suffices to count the number of solutions of the equation $u'v' = t$, $t \neq 0$. There are $q-1$ choices for u', say, and v' is completely determined. The result follows.

We conclude that

$$|g(k)|^2 = q + (q-1) \sum_{t \in \mathbb{F}_q} e^{\frac{2\pi i k t}{q}} = q.$$

Suppose that -1 is not a square in $\mathbb{F}_q$. It follows that

$$g(k) + \overline{g(k)} = \sum_{t \in \mathbb{F}_q} e^{\frac{2\pi i k t}{q}} + e^{-\frac{2\pi i k t}{q}}$$

runs over each of the elements of $\mathbb{F}_q$ exactly twice and thus equals 0. It follows that $g(k)$ is purely imaginary. When -1 is a not square in $\mathbb{F}_q$, then $\pm i$ is simply replaced by a different constant. See, for example, [**29**]. This completes the proof of Lemma 11.4.

We now prove Lemma 11.3. Keep a look out for the "counting machine" that we introduced in the previous Chapter.

$$\widehat{S}_r(m) = q^{-d} \sum_{\{x \in \mathbb{F}_q^d : |x|^2 = r\}} e^{-\frac{2\pi i x \cdot m}{q}}$$

$$= q^{-d} \sum_{x \in \mathbb{F}_q^d} q^{-1} \sum_{j \in \mathbb{F}_q} e^{\frac{2\pi i j(|x|^2 - r)}{q}} e^{-\frac{2\pi i x \cdot m}{q}}$$

$$= q^{-d-1} \sum_{j \in \mathbb{F}_q^*} e^{-\frac{2\pi i j r}{q}} \sum_{x \in \mathbb{F}_q^d} e^{\frac{2\pi i j |x|^2}{q}} e^{-\frac{2\pi i x \cdot m}{q}}$$

$$
\begin{aligned}
&= q^{-d-1} \sum_{j \in \mathbb{F}_q^*} e^{-\frac{2\pi i j r}{q}} G(-m, -j) \\
&= q^{-d-1} \sum_{j \in \mathbb{F}_q^*} e^{-\frac{2\pi i j r}{q}} (\pm i)^d q^{\frac{d}{2}} e^{-\frac{2\pi i |m|^2}{4j}} \\
&= q^{-\frac{d}{2}} q^{-1} (\pm i)^d \sum_{j \in \mathbb{F}_q^*} e^{-\frac{2\pi i}{q}(jr + \frac{|m|^2}{4j})}.
\end{aligned}
$$

This reduces the proof of Lemma 11.3 to the following Kloosterman sum estimate due to Andre Weil ([**47**]). We do not give a proof here but we encourage the reader to look one up! See, for example, [**22**] or [**31**] for an elementary and self-contained proof.

Lemma 11.6. *If q is a prime, then*

$$
\left| \sum_{j \in \mathbb{F}_q^*} e^{-\frac{2\pi i}{q}(jr + j^{-1} r')} \right| \lesssim \sqrt{q}
$$

for any $r, r' \in \mathbb{F}_q$.

We now prove that $\# S_r \approx q^{d-1}$. Using the material above,

$$
\begin{aligned}
\sum_{x \in \mathbb{F}_q^d} \left| \widehat{S}_r(x) \right|^2 &= q^{-d} q^{-2} \sum_{x \in \mathbb{F}_q^d} \sum_{u,v \in \mathbb{F}_q^*} e^{\frac{2\pi i}{q}(r(u-v) + |x|^2 (u^{-1} - v^{-1}))} \\
&= q^{-d-2} \sum_{\{(u,v) \in \mathbb{F}_q^* \times \mathbb{F}_q^* : u \neq v\}} e^{\frac{2\pi i (u-v) r}{q}} q^{\frac{d}{2}} + q^{-2} \sum_{u \in \mathbb{F}_q^*} 1 \\
&= O(q^{-1}).
\end{aligned}
$$

It follows that

$$
\# S_r = \sum_{y \in \mathbb{F}_q^d} S_r^2(x) = q^d \sum_{x \in \mathbb{F}_q^d} \left| \widehat{S}_r(x) \right|^2 = O(q^{d-1}),
$$

as desired.

We really hope that this book made you think a little bit, and that you will consider exploring this subject matter deeper. We also encourage you to reread sections of the book to see if, after some time, you can get even more out of them. Thanks for reading!

Chapter 12

Applications of the Erdős distance problem

The question often posed to us is: “Why should anyone who is not an active Erdős follower care about the Erdős distance problem?” The purpose of this chapter is to answer this question without getting too deeply into the politics of how different areas of mathematics relate to each other. We do this by giving two analytic examples designed to illustrate connections between the Erdős distance problem and some interesting problems in classical analysis and geometric measure theory. These examples were the ones that originally convinced the second listed author to study the Erdős distance problem about a decade ago.

This section assumes knowledge of basic mathematical analysis. The readers who are not familiar with the terminology and the background are encouraged to explore the theories alluded to in the following section.

A widely known mathematical fact is that the unit cube $[0,1]^d$, or a torus $\mathbb{T}^d$, depending on how one wants to look at it, possesses an orthogonal basis of exponentials. More precisely, the collection

$$\{e^{2\pi i x \cdot m} : m \in \mathbb{Z}^d\}$$

is an orthogonal basis for $L^2([0,1]^d)$. An interesting and much studied question is to determine which domains in $\mathbb{R}^d$ also possess an orthogonal bases of exponentials. More precisely, the problem is to determine, given a domain Ω, whether there exists a set $A \subset \mathbb{R}^d$ such that

$$\{e^{2\pi i x \cdot a} : a \in A\} \tag{12.1}$$

is an orthogonal basis for $L^2(\Omega)$.

An interested reader can take a look at [**13**] and the references contained therein for a description of this remarkable problem and its variants. Here we focus on a particular instance of this question, namely the question of whether the ball $B_d = \{x \in \mathbb{R}^d : x_1^2 + \cdots + x_d^2 \leq 1\}$ possesses and orthogonal basis of exponentials if $d \geq 2$. This question was raised by Fuglede in 1974, who resolved it in the case $d = 2$ using a fairly complicated analytic argument. In 1999, Nets Katz, Steen Pedersen and Alex Iosevich resolved this problem completely in [**14**], in all dimensions by reducing it to the Erdős distance problem. We now give an outline of this argument.

Assume, for the sake of contradiction, that $L^2(B_d)$ possesses an orthogonal basis of exponentials. This means that there exists $A \subset \mathbb{R}^d$ such that (12.1) holds. Orthogonality means that

$$\int_{B_d} e^{2\pi i x \cdot (a-a')} dx = 0 \tag{12.2}$$

if $a \neq a' \in A$. It follows by continuity that A is separated in the sense that there exists $c > 0$ such that $|a - a'| \geq c > 0$ for all $a \neq a' \in A$.

With a bit more work, one can show that A is well-distributed in the sense that there exists $C > c > 0$ such that every cube of side-length C contains at least one point of A. This is a straightforward analytic argument, worked out completely in [**13**], made even easier by the fact that the Fourier transform of the characteristic function of the ball has good decay properties at infinity by the method of stationary phase. See, for example, [**42**] and the references contained therein.

We now invoke (12.2) and the definition of a Bessel function (see, for example, [**42**]) to see that

$$0 = \int_{B_d} e^{2\pi i x\cdot(a-a')}dx = C|a-a'|^{-\frac{d}{2}} J_{\frac{d}{2}}(2\pi|a-a'|), \tag{12.3}$$

where J_z is the Bessel function of order z. It is well-known (see, for example, [**42**]) that zeros of Bessel functions are uniformly separated. Combining all these observations we see that if we choose an cube Q_R of side-length R, very large, in $\mathbb{R}^d$, than it contains $\approx R^d$ points of A. By (12.3) it follows that the number of distinct distances between the elements of $A \cap Q_R$ is at most $C'R$ for some $C' > 0$.

In summary, we have shown that if $L^2(B_d)$ has an orthogonal basis of exponentials, then there exists a set $S \subset \mathbb{R}^d$ with $\approx R^d$ points, such that

$$\#\Delta(S) \leq C'R.$$

As the reader shall see, it is not difficult to show, using the methods of this book, that such sets do not exist. It is conjectured that if $\#S \approx R^d$, then $\#\Delta(S) \gtrsim R^2$. While this is not known, it is fairly simply to show that $\#\Delta(S) \gtrsim R^{1+\epsilon}$ for some $\epsilon > 0$. Thus we have seen that a fairly simple application of the Erdős distance problem techniques resolve a problem in classical analysis that was open for many years.

Our second example illustrates an application of the Erdős Integer Distance Principle, discussed in detail in Chapter 2, in a similar context.

Theorem 12.1. *(Erdős Integer Distance Principle) Let $E \subset \mathbb{R}^d$ such that E is infinite and $\Delta(E) \subset \mathbb{Z}$. Then E is a subset of a line.*

A refinement of this principle was used by Misha Rudnev and the second author to prove the following result in [**17**].

Theorem 12.2. *Let $K \subset \mathbb{R}^d$, $d \geq 2$, be a convex body, symmetric about the origin, with a smooth boundary and everywhere non-vanishing curvature. Let $A \subset \mathbb{R}^d$ such that*

$$\{e^{2\pi i x\cdot a} : a \in A\}$$

is orthogonal with respect to K in the sense that

$$\int_K e^{2\pi i x\cdot(a-a')}dx = 0 \textit{ whenever } a \neq a' \in A.$$

Then the following hold:

- *If* $d \neq 1 \mod (4)$, *then A is finite.*
- *If* $d \equiv 1 \mod (4)$ *and A is infinite, then A is a subset of a line.*

An outline of the proof is the following. Using the method of stationary phase, we see that

$$\widehat{K}(\xi) = C|\xi|^{-\frac{d+1}{2}} \cos\left(2\pi\left(\rho^*(\xi) - \frac{d-1}{8}\right)\right) + O(|\xi|^{-\frac{d+3}{2}}), \tag{12.4}$$

where

$$K = \{x \in \mathbb{R}^d : \rho(x) \leq 1\},$$

with $\widehat{K}(\xi)$ representing the Fourier transform of the characteristic function of the body K, ρ denoting the Minkowski functional of K, and ρ^* is the dual functional defined by

$$\rho^*(\xi) = \sup_{x\in K} x \cdot \xi.$$

Let A be as in the statement of the theorem above. Define the ρ-distance set via

$$\Delta_\rho(A) = \{\rho^*(a-a') : a, a' \in A\}.$$

The formula (12.4) combined with the orthogonality hypothesis does not quite tell us that $\Delta_\rho(A) \subset \mathbb{Z}$, but it does tell us that $\Delta_\rho(A)$ is asymptotically close to shifted integers, which, one can show, is good enough to deduce the conclusion of the classical Erdős Integer Distance Principle from which the conclusion of Theorem 12.1 follows.

Appendix A

Hyperbolas in the plane

We will not exhaust the theory of hyperbolas here, but we will illustrate a few basic concepts so that the portion of the text concerning them can make sense. Fixing two points F_1 and F_2 in the plane, and a positive real number $a \leq |F_1 - F_2|$, a hyperbola is defined by the set

$$H_{F_1,F_2,a} = \{P \in \mathbb{R}^d : ||P - F_1| - |P - F_2|| = 2a, \} \tag{A.1}$$

where $|\cdot|$ denotes the standard Euclidean metric.

If this is confusing, another way to think of hyperbolas is as the locus of points satisfying a certain equation. Recall that one way to write the equation of a general circle or ellipse is:

$$\left(\frac{x-h}{a}\right)^2 + \left(\frac{y-k}{b}\right)^2 = r^2.$$

An analogous equation for general hyperbolas is:

$$\left(\frac{x-h}{a}\right)^2 - \left(\frac{y-k}{b}\right)^2 = r^2.$$

There are plenty of other ways to characterize hyperbolas, but this should suffice for now. Notice, if the parameters defining a hyperbola have certain values, it could actually be a single line! We call

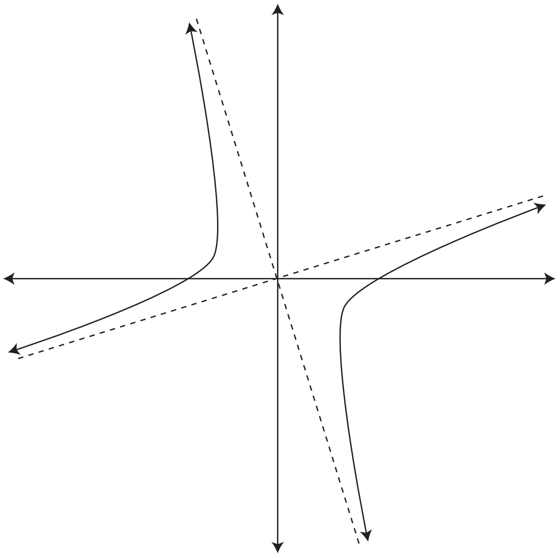


Figure A.1. This is a Cartesian plane with a single hyperbola drawn on it. The dotted lines depict the asymptotes of the hyperbola. Unlike some other conic sections, this one often has two parts.

such uninteresting hyperbolas *degenerate*. We will not discuss them further here.

Figure A.1 shows a typical hyperbola, centered at the origin. Many hyperbolas, when viewed on a large enough scale, appear to be a pair of intersecting lines. The lines that the extremities of the hyperbolas appear to behave like are called *asymptotes*. They are shown in Figure A.1 as dotted lines.

Now, the main reason that we introduced hyperbolas was to prove the Erdős integer distance principle. The property of hyperbolas that we needed was that they do not intersect each other much, given some reasonable constraints. Figure A.2 shows an example of two distinct, non-degenerate hyperbolas that intersect each other four times. In the next exercise, you will show that this is as many times as any pair of distinct, non-degenerate hyperbolas can intersect.

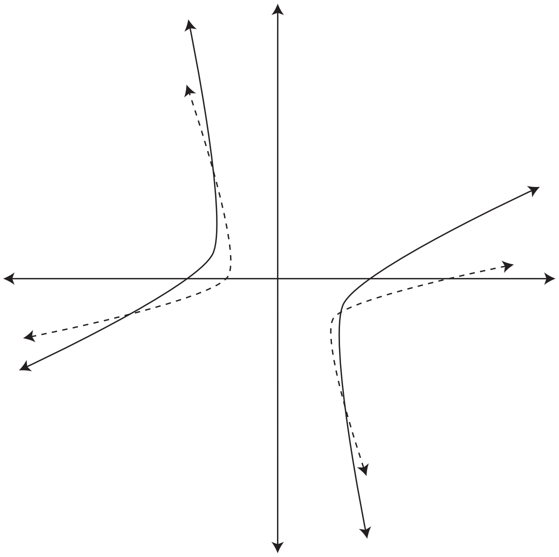


Figure A.2. This is a Cartesian plane with two hyperbolas drawn on it. This time, one hyperbola is merely drawn in normally, and the other hyperbola is drawn in with dotted lines. This is to show which half goes with which other half of each hyperbola.

Exercise A.1. Let $d = 2$. Suppose that segments F_1F_2 and $F_1'F_2'$ are not parallel. Show

$$\#(H_{F_1,F_2,a} \cap H_{F_1',F_2',a'}) \leq 4.$$

Appendix B

Basic probability theory

The title of this appendix is much too grand. We will make no effort to review probability theory in any sort of generality. Instead, we shall treat a very special case– the coin flipping experiment. The purpose of this style of presentation is to make the probabilistic argument in the book accessible to anyone, even if they have absolutely no background in probability.

Everyone knows that when you flip a coin, you have a "fifty-fifty" chance of getting heads or tails. This is because there is one outcome that corresponds to heads, and two possible outcomes total. So we can quantify this by saying that the probability of the coin landing heads up is $\frac{1}{2}$, or .5. We will write this like

$$\mathbb{P}(\text{heads}) = .5.$$

Any number between 0 and 1 can be a probability, but numbers outside of that range cannot. The *probability* of a given event, where each of the individual outcomes are equally likely, can be computed as,

$$\frac{\textit{number of possible outcomes corresponding to the given event}}{\textit{number of total possible outcomes}},$$

which is $\frac{1}{2}$ in the case of a coin landing heads up, since there is one possibility of the coin landing heads up, and two equally likely possibilities total. A *random variable* is a variable that can take certain values with some corresponding probabilities. In this case, the random variable will be the outcome of the flip. It can take the value "heads" with probability $\frac{1}{2}$, and "tails" with the same probability. Of course, the sum of all of the probabilities of all of the possible outcomes will be 1, as it is in our special case.

Since flipping a single coin is not all that interesting, we will now consider flipping several coins. The first thing to notice is that the outcome of one coinflip does not affect the outcome of another. We say that the events are *independent* of one another. An obvious contrast to this is the example of pulling cards of a given suit out of a deck of cards. If we wanted the probability of pulling a single heart out of a standard deck of playing cards is $\frac{13}{52} = \frac{1}{4}$. However, if we drew a heart out on our first try, and did not return it to the deck, as soon as we try to pull another heart out, the probability of getting a heart becomes $\frac{12}{51}$. This is because there is one fewer heart card in the deck. Of course, it is a different story altogether if the first card we pulled out was not a heart to begin with...

As you can see, the independence of the coin flips will considerably simplify the calculations of various probabilistic quantities. Now we can turn our attention to our primary object of study for this section: expectation. Since the individual events our independent, it does not matter in which order they occur, or if they happen simultaneously. So we can consider many coin flips at once and see what happens. If you flip a coin ten times, how many times do you expect to get heads? You can probably assume that your coin will land heads up five times, or half of the time. This is because you have $10 \times \mathbb{P}(\text{heads}) = 10 \times .5 = 5$. What we have just done is computed the *expected value* of the number of the number of heads. If our events are independent, the expected value of the number of a particular type of event occuring is the number of trials, (in this case coin flips) times the probability of the given outcome at each trial, (in this case, the probability of heads at each flip). To formalize this, if we flip the coin n times, the basic formula is

$$\mathbb{E}(\text{heads}) = n \cdot \mathbb{P}(\text{heads}).$$

In some situations, a given event could have several outcomes associated with it. For example, if we were flipping a coin three times, and our event was that we got two or more heads, then we would have to account for each of the three possible outcomes with two heads, plus the one possible outcome of all three heads.

To deal with more general situations, we add up the probability of each event, times the value of each event. To give an example of this, suppose you are given the chance to play a particular kind of lottery. There are two ways to win. You have a five percent chance of winning ten dollars, and a one percent chance of winning twenty dollars. So how much do you expect to win each time you play?

$$.05 \cdot 10 + .01 \cdot 20 = .70.$$

So you can expect to win seventy cents each time. Now, we can go on to speculate how much you should be willing to pay for such a game, but that is not what we are looking for today. In our applications, we will assume that every event occurs with the same probability. For an example of this sort, suppose that there are n marbles on the floor, and you pick each one up with probability p.

$$\underbrace{p \cdot 1 + p \cdot 1 + \cdots + p \cdot 1}_{n \text{ times}}$$

Then you can expect to pick up about np marbles total. Keep this example in mind as you read through the sections where expectation is used.

So as you can see, expectation can be viewed as a kind of average. We are using precise mathematical language to express ideas such as, “If you flip a coin ten times, you can expect to get five heads on average.” When dealing with large sets, it is often useful to be able to make statements about the behavior of elements in your set on average.

Another nice feature of expectation is that it is *linear*. That is to say, the expected behavior of several events is the sum of the expected behavior of the individual events. So when we consider several types of conditions, we can take expected values at any time it is convenient. This point is illustrated in Chapter 4.

Appendix C

Jensen's inequality

As we did with Cauch-Schwarz, we will prove a form of this inequality from the ground up, just by looking at seemingly uninteresting facts and drawing some interesting and useful conclusions. We will also illustrate the concept of induction. We start by defining what it means for a function to be *convex.* We call a function, f, convex if

$$f\left(\theta_1 x_1 + \theta_2 x_2\right) \leq \theta_1 f\left(x_1\right) + \theta_2 f\left(x_2\right), \tag{C.1}$$

where, θ_1 and θ_2 are positive, and $\theta_1 + \theta_2 = 1$.

We call this convex because the graph of such a function will look like the underside of a convex body. If we know that f is convex, then for any appropriate θ_1 and θ_2, we can be assured that (C.1) holds.

Since it holds for two pairs of x's and θ's, we could try to show that it holds for three pairs of x's and θ's, then four, and so on. However, at some point, we would have to stop, as our lives are only so long! To address this, there is a process called *induction*, by which we can derive statements for as many pairs of x's and θ's as we wish.

In general, if you have a statement that you want to show is true for any number, you start by showing that it holds for some small value of n. This is called the *base case.* Then you assume that it holds for some arbitrary value of n and try to show that this implies that the statement is true for $n+1$. Since you have shown that it is true for the base case, and that one implies the next, you know that it is true for the next case after the base case, and the case after that, etc... This is not the only way that induction works, but it is the simplest, and most often employed.

In our scenario, the base case will be $n=2$. Consider (C.1) to be our desired statement for two pairs. Since we have already shown our statement to be true for $n=2$, we can proceed by trying to show that validity for n implies validity for $n+1$. So assume that something like (C.1) holds for n numbers, $x_1, x_2, ..., x_n$, where the sum of the θ_i's is 1.

$$f\left(\sum_{i=1}^{n} \theta_i x_i\right) \leq \sum_{i=1}^{n} \theta_i f\left(x_i\right). \tag{C.2}$$

Then we will use this to show that it holds for $n+1$ numbers. Our final goal will be to show:

$$f\left(\sum_{i=1}^{n+1} \theta_i x_i\right) \leq \sum_{i=1}^{n+1} \theta_i f\left(x_i\right)$$

The idea is to write the left hand side for $n+1$ numbers, and use $(C.1)$ and $(C.2)$ to get an appropriate right hand side. So, if we want to get something that has two terms in the argument, or input, of the function, like $(C.1)$, we should seperate the sum somehow. We also want to use $(C.2)$, so we should have a sum of n numbers somewhere. Keeping these goals in mind, one logical approach would be the following:

$$f\left(\sum_{i=1}^{n+1} \theta_i x_i\right) = f\left(\theta_1 x_1 + \sum_{i=2}^{n+1} \theta_i x_i\right),$$

Now, we have two terms inside the argument of our function. However, we still do not quite have them in the form we want yet, as the second term does not have a "θ"-like factor in front of it. So we will put one in. We need a number, such that adding it to θ_1 will give us 1. So we need the factor $(1-\theta_1)$. Now, we are not allowed to just throw it in front of the sum, but we can multiply and divide by it. This will yield:

$$f\left(\theta_1 x_1 + \sum_{i=2}^{n+1} \theta_i x_i\right) = f\left(\theta_1 x_1 + (1-\theta_1) \sum_{i=2}^{n+1} \frac{\theta_i}{(1-\theta_1)} x_i\right).$$

If we view the sum as one big number, call it x_2', and the $(1-\theta_1)$ as the factor in front of it, θ_2', then we can use the convexity of f:

$$f\left(\theta_1 x_1 + \theta_2' x_2'\right) \leq \theta_1 f\left(x_1\right) + \theta_2' f\left(x_2'\right).$$

Substitute everything back in, and we have

$$\theta_1 f\left(x_1\right) + (1-\theta_1) f\left(\sum_{i=2}^{n+1} \frac{\theta_i}{(1-\theta_1)} x_i\right).$$

We are almost done! Now we have to deal with the sum of the remaining n numbers in the sum. We want to employ the n number inequality to the sum term. The sum of the remaining θ_i's is $(1-\theta_1)$. So the sum of the last n of the $\frac{\theta_i}{1-\theta_1}$'s is 1. This means that we can use $(C.2)$ on the latter sum.

$$\theta_1 f\left(x_1\right)+(1-\theta_1) f\left(\sum_{i=2}^{n+1} \frac{\theta_i}{(1-\theta_1)} x_i\right) \leq \theta_1 f\left(x_1\right)+(1-\theta_1) \sum_{i=2}^{n+1} \frac{\theta_i}{(1-\theta_1)} f\left(x_i\right)$$

We have just shown one form of Jensen's inequality.

Theorem C.1. *Given a convex function f, and a sequence of n positive numbers, $\{x_i\}_{i=1}^n$,*

$$f\left(\frac{\sum_{i=1}^n x_i}{n}\right) \leq \frac{\sum_{i=1}^n f\left(x_i\right)}{n}.$$

Exercise C.1. Use induction to show that

$$1 + 2 + ... + n = \frac{n(n+1)}{2}.$$

Show that it is true for $n = 1$, then assume that it is true for n, and show that that implies that it is true for $n + 1$.

Exercise C.2. Find the conditions that allow us to interpret Jensen's inequality as

$$f\left(\mathbb{E}(x)\right) \leq \mathbb{E}\left(f(x)\right).$$

Bibliography

[1] P. K. Agarwal, E. Nevo, J. Pach, R. Pinchasi, M. Sharir, and S. Smorodinsky, *Lenses in arrangements of pseudo-circles and their applications*, J. ACM, to appear.

[2] M. Ajtai, V. Chvatal, M. Newborn, and E. Szemerédi, *Crossing-free subgraphs*, Ann. Discrete Mathematics **12** (1982) 9–12.

Cutting circles into pseudo-segments and improved bounds for incidences, Discrete Comput. Geom. **28** (2002), no. 4, 475–490.

Falconer conjecture, spherical averages and discrete analogs, Towards a theory of geometric graphs, 15–24, Contemp. Math., 342, Amer. Math. Soc., Providence, RI, 2004.

[3] J. Beck, *On the lattice property of the plane and some problems of Dirac, Motzkin and Erdős in combinatorial geometry*, Combinatorica **3** (1983), no. 3-4, 281–297.

Unit distances, J. Combinatorical Theory A **37** (1984), no. 3, 231–238.

[4] J. Bourgain, N. Katz, and T. Tao *A sum-product estimate in finite fields, and applications* Geom. Funct. Anal. **14** (2004), 27-57.

Hausdorff dimension and distance sets, Israel J. Math. **87** (1994), no. 1-3, 193–201.

On the Erdős-Volkmann and Katz-Tao ring conjectures, Geom. Funct. Anal. **13** (2003), no. 2, 334–365.

[5] P. Brass, *Erdős distance problems in normed spaces*, Discrete Comput. Geom. **17** (1997), no. 1, 111–117.

[6] F. R. K. Chung, *The number of different distances determined by n points in the plane*, J. Combin. Theory Ser. A **36** (1984), no. 3, 342–354.

[7] F. R. K. Chung, E. Szemerédi and W. T. Trotter, *The number of different distances determined by a set of points in the Euclidean plane*, Discrete Comput. Geom. **7** (1992), no. 1, 1–11.

[8] K. L. Clarkson, H. Edelsbrunner, L. Guibas, M. Sharir, and E. Welzl, *Combinatorial complexity bounds for arrangements of curves and spheres*, Discrete Comput. Geom. **5** (1990), no. 2, 99–160.

[9] P. Erdős, *On sets of distances of n points*, Amer. Math. Monthly **53** (1946) 248–250.

[10] P. Erdős, *Integral distances*, Bull. Amer. Math. Soc. **51** (1945).

[11] J. S. Garibaldi, *A Lower Bound for the Erdős Distance Problem for Convex Metrics*, preprint.

[12] A. Iosevich, *Curvature, combinatorics and the Fourier transform*, Notices Amer. Math. Soc. **48** (2001), no. 6, 577–583.

[13] A. Iosevich, *Fourier analysis and geometric combinatorics*, World Scientific volume dedicated to the annual Padova lectures in analysis (2008).

[14] A. Iosevich, N. Katz and S. Pedersen, *Fourier bases and a distance problem of Erdos*, Math Research Letters, **6**, Number 2, (1999), Number 2, 105-128.

[15] A. Iosevich and I. Łaba, *Distance sets of well-distributed planar sets*, Discrete Comput. Geom. **31** (2004), no. 2, 243–250.

[16] A. Iosevich and I. Łaba, *K-distance, Falconer conjecture, and discrete analogs*, preprint.

[17] A. Iosevich and M. Rudnev, *A combinatorial approach to orthogonal exponentials*, Int. Math. Res. Notices, Volume 2003, Number 50, Pp. 2671-2685.

[18] A. Iosevich and, M. Rudnev, *Erdős distance problem in vector spaces over finite fields*, Transactions of the American Mathematical Society, (2007).

[19] D. Hart A. Iosevich D. Koh and M. Rudnev, *Averages over hyperplanes, sum-product theory in finite fields, and the Erdos-Falconer distance conjecture*, (accepted for publication by Transactions of the AMS).

[20] A. Iosevich N. H. Katz, and T. Tao, *Convex bodies with a point of curvature do not have Fourier bases*, American Journal of Mathematics **123**, (2001), 115-120.

[21] A. Iosevich, A View from the Top, AMS, Providence, RI, 2007.

[22] H. Iwaniec and E. Kowalski *Analytc Number Theory* Colloquium Publications, **53** (2004).

[23] Amnon Katz, *Principles of Statistical Mechanics: The Information Theory Approach* Freeman and Company (1967).

[24] N. H. 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.

[25] N. H. Katz and G. Tardos, *A new entropy inequality for the Erdős distance problem*, Towards a Theory of Geometric Graphs, (ed. J Pach) Contempory Mathematics **342** (2004), 119–12.

[26] A. I. Khinchin, *Mathematical Foundations of Information Theory*, Dover (1957).

[27] A. G. Khovanskiĭ, *A class of systems of transcendental equations*, Dokl. Akad. Nauk SSSR **255** (1980), no. 4, 804–807.

[28] A. G. Khovanskiĭ, *Fewnomials*, Translated from the Russian by Smilka Zdravkovska, Amer. Math. Soc., Providence, RI, 1991.

[29] E. Landau *Vorlesungen ber Zahlentheorie* Chelsea Publishing Co., New York (1969).

[30] T. Leighton, *Complexity Issues in VLSI, Foundations of Computer Series*, MIT Press, Cambridge, MA, 1983.

[31] R. Lidl and H. Niederreiter, *Finite fields*, Cambridge University Press (1997).

[32] L. Ma, *Bisectors and Voronoi diagrams for convex distance functions*, dissertation (unpublished).

[33] L. Moser, *On the different distances determined by n points*, Amer. Math. Monthly **59** (1952), 85–91.

[34] M. Nathanson, *Additive number theory: inverse problems and the geometry of sumsets*, Springer, New York, NY, 1996

[35] J. Pach and P.K. Agarwal, *Combinatorial Geometry*, Wiley, New York, NY, 1995.

[36] J. Pach and M. Sharir, *On the number of incidences of points and curves*, Combin. Probab. Comput. **7** (1998), no. 1, 121–127.

[37] Imre Ruzsa, *A Problem on Restricted Sumsets* Towards a theory of geometric graphs, 245-248, Contemp. Math., 342, Amer. Math. Soc., Providence, RI, (2004).

[38] J. Solymosi and C. Tóth, *Distinct distances in the plane*, Discrete Comput. Geom. **25** (2001), no. 4, 629–634.

[39] J. Solymosi, *Note on integral distances*, Discrete Comput. Geom. **30** (2003), no. 2, 337–342.

[40] J. Solymosi, G. Tardos and C. D. Tóth, *The k most frequent distances in the plane*, Discrete Comput. Geom. **28** (2002), no. 4, 639–648.

[41] Swanepoel, Konrad, *Cardinalities of k-distance sets in Minkowski spaces*, Discrete Mathematics **197/198** (1999) 759-767.

[42] Stein, Elias M., *Harmonic Analysis* Princeton University Press, (1993).

[43] Székely, L. A. (1987) *Inclusion-exclusion formulae without higher terms*, Ars Combinatoria **23B** 7-20.

[44] L.A. Székely, *Crossing numbers and hard Erdős problems in discrete geometry*, Combin. Probab. Comput. **6** (1997), no. 3, 353–358.

[45] E. Szemerédi and W. T. Trotter, Jr., *Extremal problems in discrete geometry*, Combinatorica **3** (1983), no. 3-4,381–392.

[46] G. Tardos, *On distinct sums and distinct distances*, Adv. Math. **180** (2003), no. 1, 275–289.

[47] A. Weil, *On some exponential sums* Proc. Nat. Acad. Sci. U.S.A. **34** (1948), 204-207.