A view from the top: analysis, combinatorics and number theory

Alex Iosevich, Professor of Mathematics,
University of Missouri-Columbia

Biographical information

The 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 appears to have settled down at the University of Missouri where this book was entirely written.

Contents

Thanks

This book would not have been possible without the influence and assistance of numerous people. I wish to thank my family who suffered a certain amount of divided attention. Much thanks is due to my students and colleagues whose comments and suggestions had a profound impact on this work. It is impossible to mention everybody, but I am especially grateful to Gina Belarde, David Covert, Michael Gramlich, Zara Girnius, Lacy Hardcastle, Derrick Hart, Steve Hofmann, Tyler Jones-Salisbury, Nets Katz, James Knapp, Doowon Koh, Bill McClain, David Meyer, Lorentz Morrow, Heather Rosenblatt, Shannon Reed, Mark Rudelson, Misha Rudnev, Steve Senger, Terry Tao, Ignacio Uriarte-Tuero, Marisa Zymonopoulou and the anonymous referees for thorough proofreading and excellent recommendations.

The author is especially indebted to Shannon Lee Reed who suggested and encouraged the creation of this book during and after she enrolled in the summer program on the Besicovitch-Kakeya conjecture in August 2004 and in the capstone course I taught at the University of Missouri in the Fall of 2004. I am most grateful to her for helping me rediscover my passion for writing, which has had a profound impact on my approach to teaching and research.

I wish to thank my colleague, Loukas Grafakos, for an excellent suggestion regarding the presentation of the fact that trigonometric polynomials are dense in the space of square integrable functions on

the torus, a remark that made the presentation of the last chapter of the book much more coherent. I also wish to thank my colleague Fritz Gezstesy for making me aware of the beautiful book, entitled "*The Cauchy-Schwartz Master Class*", by J. Michael Steele ([**16**]), which significantly affected my thinking about several topics covered in this book.

Much gratitude is due to Nets Katz and Wilhelm Schlag who encouraged me to explore connections between combinatorics and harmonic analysis early in my career. Such connections have since then become a centerpiece of my research program.

I feel an infinite amount of gratitude to my parents, Michael and Svetlana Iosevich, my brother Sam, and my deceased grandparents, without whose support nothing meaningful in my life would have been possible.

I am deeply indebted to Eugene Yampolsky, a graduate student in mathematics at the University of Missouri-Columbia, who prepared many useful and attractive diagrams that are displayed in this book.

Foreword

The idea for writing this book arose as a result of conversations I had with Shannon Reed when she was a student in Math 4980, the problem solving class, at the University of Missouri during the Fall semester of 2004. This was a capstone course for upper division students designed to bring together techniques and ideas from the standard undergraduate curriculum. Another purpose for the course was to teach students problem solving techniques and to prepare them for the national Putnam examination in December. I was excited to discover over the years that these goals, far from being incompatible, work hand in hand to enhance the students' hands on knowledge of mathematics and to provide them with a glimpse into the world of research and discovery.

The idea of a capstone class is absolutely wonderful and it keeps growing on me each year. Too many undergraduate students form an impression of mathematics as being compartmentalized into specific subject areas like analysis, algebra, number theory, topology and others. The idea of unity of mathematics and interactions between fundamental areas are seldom mentioned and almost never taught. The purpose of this book is to attempt to break through this barrier. It is a hopeless task to illustrate even a small sliver of exciting and surprising connections that mathematics is so full of. The author confines himself to the connections that are near and dear to his

heart—basic analytic inequalities, probabilistic reasoning, and their connections with geometric combinatorics and number theory.

We begin the book by introducing the reader to the Cauchy-Schwartz and Hölder inequalities. Instead of continuing on to the endless, albeit interesting, world of inequalities, we immediately pursue applications to geometric problems. We hope that the natural appeal and beauty of these connections will help us make the case that, far from being solely an exercise in symbol manipulation, Cauchy-Schwartz and other dry looking estimates reflect fundamental physical realities that can be appreciated on many levels. For example, we show that the Cauchy-Schwartz inequalities can be used to estimate the number of incidences of points and lines, and sizes of projections of discrete point sets. In the course of discussing projections, we quietly sneak in the notion of interpolation which is so fundamental and unavoidable in research harmonic analysis. When presented for its own sake, this concept can appear dry and specialized. In the context of a concrete problem, however, it is instead at worst a necessary evil needed to resolve the problem at hand. These ideas and their variants occupy the first four chapters of the book.

In chapters 5-8, we move on to the finite field setting, explained in detail without any need for prerequisites, thus simplifying the calculations and eliminating the need for much formalism. This allows us to present much of what is known on the Besicovitch-Kakeya conjecture, one of the most important and central problems of modern harmonic analysis which connects the size of a set with the number of "line segments" of different "slopes" contained within. Chapters 9 and 10 are dedicated to problems and ideas that require basic counting and probabilistic reasoning, which we then connect with some interesting questions in the theory of numbers, thus putting a different perspective on calculations and concepts introduced earlier in the manuscript. Chapters 11 and 12 of the book are dedicated to trigonometric sums, and sums and integrals with applications to problems in geometry and number theory.

We do not aim for the slickest proofs or even the most elegant presentation. The idea is to get the reader to become excited about research mathematics by observing the process in which ideas evolve.

In several chapters we develop the necessary tools in the process of investigation without even alluding to the fact that these are standard results in various areas of mathematics. We aim to get across to the reader that mathematics is not discovered by reproduction and slight modification of techniques found in textbooks, but rather through a painful and often comical process of discovery and rediscovery.

While not all the problems in this book are at the cutting edge of modern mathematics, the techniques are selected precisely for their importance and ubiquity in mathematical research. Connections between different techniques and areas of mathematics are emphasized throughout and constitute one of the most important lessons this book attempts to impart.

The student is expected to work hard while reading this book. This is not bed time reading, nor is it a fantasy novel. You must have a pen and plenty of paper handy, and expect to fill up about ten pages of calculations for every page you read. Mathematics is not a spectator sport, so create in addition to reading and computing. Every time you see a theorem or a calculation, try to formulate a new one. Every time you see a proof, try to find a better one. And most importantly, have fun!

My goal is to make this book interesting and accessible to anyone with the basic knowledge of high school mathematics who is curious about research mathematics. Several chapters require knowledge of calculus of several variables, and this is clearly indicated in the beginning of each of those chapters.

On the other hand, many topics of this book may even be of interest to graduate students in mathematics and professional researchers. While the vast majority of techniques described in this book are well known to professional mathematicians, the perspective and interlacing of topics and ideas may turn out to be unusual and even surprising.

Chapter 1

The Cauchy-Schwarz inequality

In this section we shall follow a procedure often considered nasty, but the one I hope to convince you to appreciate. We shall work backwards, discovering concepts as we go along, instead of stating them ahead of time. No background beyond high school mathematics is required to read this chapter. For further information on the material presented here, the reader is encouraged to take a look at J. Michael Steele's beautiful book, entitled "*The Cauchy-Schwarz Master Class*" ([**16**]).

A quick perspective before we begin the nuts and bolts of the mathematical discussion. Inequalities are a dime a dozen. The statement that $2 \leq 3$ is a true inequality, but it is meaningless. An interesting inequality is one that comes close to not being true, but does not quite cross that precarious threshold. This is not much different than any area of learning. Saying that the United States has a bigger per capita income than Chad is a true but meaningless statement. On the other hand, the fact that the People's Republic of Congo has land area equal to nearly two thirds of the land area of the United States is much more precise, interesting and surprising. To put it bluntly, to say something interesting, one must walk on the very edge of the

cliff of falsehood, yet never fall off. It is time to begin the perilous journey.

Let a and b denote two real numbers. Then

$$(a-b)^2 \geq 0.$$

This statement is so vacuous, you are probably wondering why I am telling you this. Nevertheless, expand the left hand to see that

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

which implies that

$$ab \leq \frac{a^2+b^2}{2}. \tag{1.1}$$

Now consider

$$A_N = \sum_{k=1}^{N} a_k = a_1 + \cdots + a_N,$$

$$B_N = \sum_{k=1}^{N} b_k = b_1 + \cdots + b_N,$$

where $a_1, \ldots, a_N$, and $b_1, \ldots, b_N$ are real numbers. Let

$$X_N = \left(\sum_{k=1}^{N} a_k^2 \right)^{\frac{1}{2}},$$

$$Y_N = \left(\sum_{k=1}^{N} b_k^2 \right)^{\frac{1}{2}}.$$

Our goal is to take advantage of (1.1). Let's take a look at

$$\begin{aligned} \sum_{k=1}^{N} a_k b_k &= X_N Y_N \sum_{k=1}^{N} \frac{a_k}{X_N} \cdot \frac{b_k}{Y_N} \\ &\leq X_N Y_N \sum_{k=1}^{N} \left[\frac{1}{2} \left(\frac{a_k}{X_N} \right)^2 + \frac{1}{2} \left(\frac{b_k}{Y_N} \right)^2 \right]. \end{aligned} \tag{1.2}$$

Exercise 1.1. Explain why if C is a constant, then

$$\sum_{k=1}^{N} C a_k = C \sum_{k=1}^{N} a_k.$$

Exercise 1.2. 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 Exercise 1.1 and 1.2 to rewrite the right hand side of (1.2) in the form

$$\begin{aligned} & X_N Y_N \frac{1}{2}\frac{1}{X_N^2}\sum_{k=1} a_k^2 + X_N Y_N \frac{1}{2}\frac{1}{Y_N^2}\sum_{k=1}^{N} b_k^2 \\ &= X_N Y_N \frac{1}{2}\frac{1}{X_N^2} X_N^2 + X_N Y_N \frac{1}{2}\frac{1}{Y_N^2} Y_N^2 \\ &= \frac{1}{2} X_N Y_N + \frac{1}{2} X_N Y_N = X_N Y_N. \end{aligned} \tag{1.3}$$

Using (1.3) and putting everything together, we obtain the Cauchy-Schwarz inequality:

Theorem 1.4. *Let a_k, b_k be real numbers. Then*

$$\sum_{k=1}^{N} a_k b_k \leq \left(\sum_{k=1}^{N} a_k^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^{N} b_k^2\right)^{\frac{1}{2}}. \tag{1.5}$$

Exercise 1.3. Prove that that equality in (1.5) occurs if and only if $a_k = b_k$ for all k. Hint: How did this all begin? Surely the equality in the inequality $(a-b)^2 \geq 0$ happens if and only if $a = b$...

We now use a variant of the same procedure to deduce the following generalization of the Cauchy-Schwarz inequality known as Hölder's inequality.

Theorem 1.6. *Let $1 < p < \infty$ and define the "dual" exponent p' by the equation*

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

Then

$$\sum_{k=1}^{N} a_k b_k \leq \left(\sum_{k=1}^{N} |a_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{N} |b_k|^{p'}\right)^{\frac{1}{p'}}. \tag{1.7}$$

Following the proof of Cauchy-Schwartz above it is not difficult to see that it suffices to prove that

$$ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'}. \tag{1.8}$$

Exercise 1.4. Check the details of this and demonstrate line by line that Hölder's inequaity indeed follows from (1.8)!

To prove (1.8) we need to use first year calculus. If you are not familiar with calculus, take (1.8) for granted and move for now. Do not worry...

If you are familiar with basic calculus, set $a^p = e^x$ and $b^{p'} = e^y$. 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 $t \in [0, 1]$,

$$e^{(1-t)x+ty} \leq (1-t)e^x + te^y.$$

Let

$$F(t) = (1-t)e^x + te^y - e^{(1-t)x+ty}$$

and observe that $F(0) = F(1) = 0$. Also observe that

$$F''(t) = -(x-y)^2 e^{(1-t)x+ty} \leq 0.$$

What do we have? A twice differentiable function F is equal to 0 at $t = 0$ and $t = 1$. Moreover, F is concave up because $F''(t) \leq 0$. We are forced to conclude that F is always non-negative and (1.8) follows.

Exercise 1.5. Prove the Cauchy-Schwartz inequality in the following way. Let $a = (a_1, \ldots, a_N)$ and $b = (b_1, \ldots, b_N)$. Define

$$< a, b >= a_1 b_1 + \cdots + a_N b_N,$$

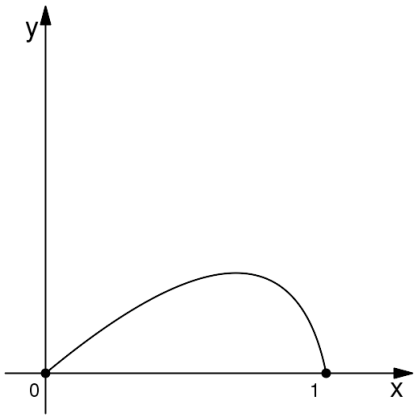
and define

$$||a||^2 = a_1^2 + a_2^2 + \cdots + a_N^2.$$

Cauchy-Schwarts takes the form

$$< a, b > \leq ||a|| \cdot ||b||.$$

Consider $< a - tb, a - tb >$, expand it out, write it as a quadratic polynomial in t, minimize it in t, and complete the proof.



Exercise 1.6. Prove the following related inequality. Let $x_1, \ldots, x_d$ denote positive real numbers. Then

$$(x_1 \cdot x_2 \cdot \cdots \cdot x_n)^{\frac{1}{n}} \leq \frac{x_1 + x_2 + \cdots + x_n}{n}. \tag{1.9}$$

Hint: There are many ways to do this, but I suggest exploring the following idea. Prove this inequality for $n = 2$ and then extend it, using induction, to $n = 2^k$, $k = 1, 2 \ldots$. Then prove that if the result holds for $n + 1$, then it must hold for n, thus filling the gaps between the powers of two.

Another approach: Write $a_j = e^{\log(a_j)}$ and use the convexity of the exponential function, i.e, the fact that

$$e^{t_1 x_1 + t_2 x_2 + \cdots + t_n x_n} \leq t_1 e^{x_1} + t_2 e^{x_2} + \cdots + t_n e^{x_n}$$

with $t_j \geq 0$ and

$$t_1 + t_2 + \cdots + t_n = 1.$$

Exercise 1.7. Let $x_1, \ldots, x_n$ and $a_1, \ldots, a_n$ be positive real numbers. Then

$$x_1^{a_1} \cdot x_2^{a_2} \cdot \cdots \cdot x_n^{a_n} \leq \frac{(x_1 + \cdots + x_n)^{a_1 + \cdots + a_n}}{a_1 \cdot a_2 \cdot \cdots \cdot a_n} a_1^{a_1} \cdot a_2^{a_2} \cdot \cdots \cdot a_n^{a_n}.$$

Is it possible to choose a_js in a way that reduces this inequality to either (1.7) or (1.9)? If the answer is yes, demonstrate it. If the answer is no, explain why not? Hint: Under what conditions does the equality hold in all the inequalities involved in this exercise.

Exercise 1.8. Generalize (1.8) to see that

$$a_1 a_2 \dots a_n \leq \frac{a_1^{p_1}}{p_1} + \dots + \frac{a_n^{p_n}}{p_n},$$

where each $p_j > 1$ and

$$\frac{1}{p_1} + \frac{1}{p_2} + \dots + \frac{1}{p_n} = 1.$$

While this follows fairly easily from the method of the previous exercise, solve this proof using Lagrange multipliers. More precisely, let

$$f(a_1, \dots, a_n) = \frac{a_1^{p_1}}{p_1} + \dots + \frac{a_n^{p_n}}{p_n},$$

and let

$$g(a_1, \dots, a_n) = a_1 a_2 \dots a_n.$$

Now minimize $f(a_1, \dots, a_n)$ under the constraint $g(a_1, \dots, a_n) = c$, where c is an arbitrary positive number.

1. Notes, remarks and difficult questions

Many high school students have seen the Cauchy-Schwartz inequality, at least in the case $N = 2$, without ever realizing it. Let $x = (x_1, x_2)$ and $y = (y_1, y_2)$ be two vectors in the plane. We know from high school mathematics that

$$x \cdot y = x_1 y_1 + x_2 y_2 = |x| \cdot |y| \cdot \cos(\theta),$$

where

$$|x| = \sqrt{x_1^2 + x_2^2},$$

and θ is the angle between x and y. Since $|\cos(\theta)| \leq 1$, we see that

$$|x \cdot y| \leq |x| \cdot |y|,$$

which is the Cauchy-Schwartz inequality in the case $N = 2$.

Can you use the fact that two vectors determine a plane, even in d-dimensions, to extend the above reasoning to $\mathbb{R}^d$, thus giving yet another proof of the Cauchy-Schwartz inequality?

Provided that you were able to answer the question in the previous paragraph, we now have three proofs of the Cauchy-Schwartz inequality. Do the first two inequalities rely on the assumption that the left hand side of the Cauchy-Schwartz inequality, $\sum_{k=1}^{N} a_k b_k$, is a finite sum? What about the third inequality?

Chapter 2

Projections in $\mathbb{R}^3$—the elephant makes an appearance!

In this chapter we use the Cauchy-Schwarz inequality to obtain bounds on the relationship between the number of elements in a finite set in three dimensions and the size of its projections onto coordinate two-dimensional planes. There are many boring and technical ways to describe what is going on here, but let us not do that just yet. Suppose that you have just inherited a very massive elephant from your uncle, yet you do not own a scale large enough to measure his weight. You know that the weight of the elephant is related to his volume in a fairly predictable way, but how do you measure his volume ? Take a large bed sheet and attach it right behind the elephant. Take another bed sheet and attach it directly to the side of the elephant. Finally, put a third bed sheet under the elephant. Now take a water cannon, spray the elephant from the front and see how much water has landed on the bed sheet. Repeat this procedure from the side and then from the top. What is the idea here? If the elephant has large volume, one of the bed sheets will be much more dry than the others! I mean, the elephant could have a large volume by being very tall, yet thin. This means that the bed sheet on the bottom of the elephant will be very wet as most of the water will get through. On the other hand, the bed

sheet on the side of the elephant will hardly suffer any damage at all. The purpose of this chapter is to make these silly looking common sense observations quantitative and this is where our story now takes us.

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

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

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

$$S_N = \{(k, l, 0) : k, l \text{ integers } 1 \le k \le \sqrt{N}, 1 \le l \le \sqrt{N}\},$$

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

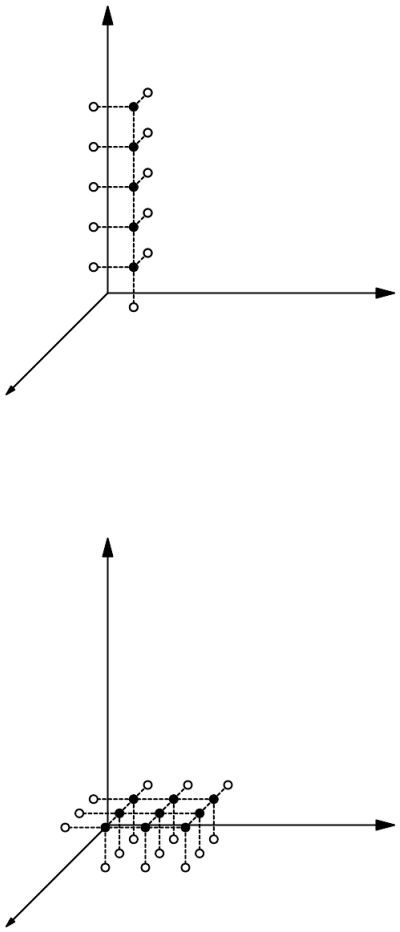
$$\pi_3(S_N) = \{(k, l) : k, l \text{ integers } 1 \le k \le \sqrt{N}, 1 \le l \le \sqrt{N}\}.$$

It follows that $\#\pi_3(S_N) = N$. On the other hand,

$$\pi_2(S_N) = \{(k, 0) : k \text{ integer } 1 \le k \le \sqrt{N}\},$$

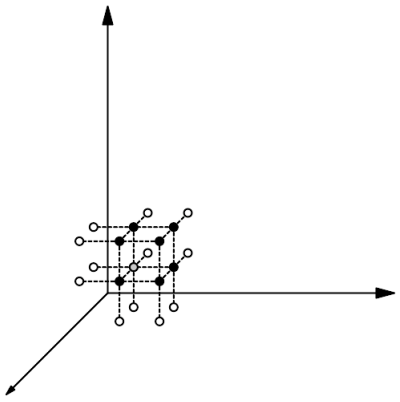
and

$$\pi_1(S_N) = \{(l, 0) : l \text{ integer } 1 \le l \le \sqrt{N}\},$$



both containing $\sqrt{N}$ elements. Again we see that it is difficult for all 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

$$S_N = \{(k,l,m) : k,l,m \text{ integers } 1 \le k,l,m \le N^{\frac{1}{3}}\},$$

where $N^{\frac{1}{3}}$ is an integer. Again, $\#S_N = N$, as required. The projections this time all look the same. We have

$$\pi_1(S_N) = \{(l,m) : l,m \text{ integers } 1 \le l,m \le N^{\frac{1}{3}}\},$$

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 be of size at least $N^{\frac{2}{3}}$. We will see here and later in these notes that C-S inequality is very useful in showing that the "symmetric" case is the "optimal", whatever that means in a given instance.

2.1. The two-dimensional case. Before starting a more detailed investigation, consider what happens in the plane in hopes of getting some ideas. Take a set of N points in $\mathbb{R}^2$ and consider projections onto the x_1-axis and the x_2-axis, respectively. Can we prove by a simple geometric argument that one of these projections must contain at least $C\sqrt{N}$ points? Well, let S_N be the set in question. Let $\chi_S(x) = 1$,

if $x \in S$, and 0 otherwise. Observe that

$$\chi_{S_N}(x_1, x_2) \leq \chi_{\pi_1(S_N)}(x_2) \cdot \chi_{\pi_2(S_N)}(x_1)$$

(see Exercise 2.5 below). It follows that

$$\begin{aligned}\sum_{x_1,x_2} \chi_{S_N}(x_1, x_2) &\leq \sum_{x_1,x_2} \chi_{\pi_1(S_N)}(x_2) \cdot \chi_{\pi_2(S_N)}(x_1) \\ &= \sum_{x_1} \chi_{\pi_2(S_N)}(x_1) \cdot \sum_{x_2} \chi_{\pi_1(S_N)}(x_2)\end{aligned}$$

(why?). It follows that

$$N = \#S_N \leq \#\pi_1(S_N) \cdot \#\pi_2(S_N) \leq \left(\max_{j=1,2} \#\pi_j(S_N)\right)^2.$$

We conclude that indeed $\max_{j=1,2} \pi_j(S_N) \geq \sqrt{N}$ as promised.

The point of considering the two-dimensional case is that while it does not entail any of the interesting complexities of the higher dimensional situation, it is based on the same intuition. Indeed, let us think about the two-dimensional case from a slightly different point of view. Suppose for a moment that $\#\pi_1(S_N)$ is smaller than $\sqrt{N}$. Then S_N must consist of at most $\sqrt{N}$ rows, by definition of π_1. On the other hand, the total number of points in all of those rows is N, by assumption. It follows that one of these rows has more than $\frac{N}{\sqrt{N}}$ points. We conclude that either $\#\pi_1(S_N) \geq \sqrt{N}$, or $\#\pi_2(S_N) \geq \sqrt{N}$, since the latter is precisely what it means for a row to have more than $\sqrt{N}$ points. This gives an "alternate" argument for the two-dimensional case. Observe that the argument given in the previous paragraph is at least superficially mechanical, involving symbolic manipulation, while the argument we just went over is visual and conceptual. Are the arguments really different, however? As an informal exercise, cut through the mechanical non-sense of the first argument and explain why it is the same as the second one.

2.2. We are ready for 3-D!. The three dimensional case is not going to fall quite so easily. To see this, let us try to run the argument of the previous paragraph. Suppose that $\#\pi_1(S_N) < N^{\frac{2}{3}}$. This means that S_N consists of fewer than $N^{\frac{2}{3}}$ columns of points over the (x_2, x_3)-plane. Since the total number of points is N, this tells us that one of the columns has more than $N^{\frac{1}{3}}$ points. This is not enough, however,

and more careful analysis is needed. The proof can be completed this way with some work and I urge you to try! We will take a slightly different road below in order to illustrate what a beautiful bookkeeping tool the C-S inequality often is.

To start our analysis of the three-dimensional case we need the following basic definition. Let S be any set. As before, define $\chi_S(x) = 1$ if $x \in S$ and 0 otherwise.

Exercise 2.1. Let S_N be as above. Then

$$\chi_{S_N}(x) \leq \chi_{\pi_1(S_N)}(x_2, x_3)\chi_{\pi_2(S_N)}(x_1, x_3)\chi_{\pi_3(S_N)}(x_1, x_2). \tag{2.1}$$

Can you give an example of a set S_N, where the left hand side is strictly smaller than the right hand side for some values of x? All values of x? For which types of sets do we always have an equality?

With Exercise 2.1 in tow, we write

$$N = \#S_N = \sum_x \chi_{S_N}(x)$$

$$\leq \sum_x \chi_{\pi_1(S_N)}(x_2, x_3)\chi_{\pi_2(S_N)}(x_1, x_3)\chi_{\pi_3(S_N)}(x_1, x_2)$$

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

$$\leq \left(\sum_{x_1,x_2} \chi^2_{\pi_3(S_N)}(x_1, x_2) \right)^{\frac{1}{2}}$$

$$\times \left(\sum_{x_1,x_2} \left(\sum_{x_3} \chi_{\pi_1(S_N)}(x_2, x_3)\chi_{\pi_2(S_N)}(x_1, x_3) \right)^2 \right)^{\frac{1}{2}}$$

$$= I \times II.$$

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

$$\#S_N \leq \sqrt{\#\pi_1(S_N)}\sqrt{\#\pi_2(S_N)}\sqrt{\#\pi_3(S_N)}. \tag{2.2}$$

Exercise 2.2. Verify each step above. Where was the Cauchy-Schwartz 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 the argument above that

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

As a result we have the following attractive result.

Theorem 2.3. *Let S_N be a subset of $\mathbb{R}^3$ consisting of N points. Then (2.2) holds and, consequently,*

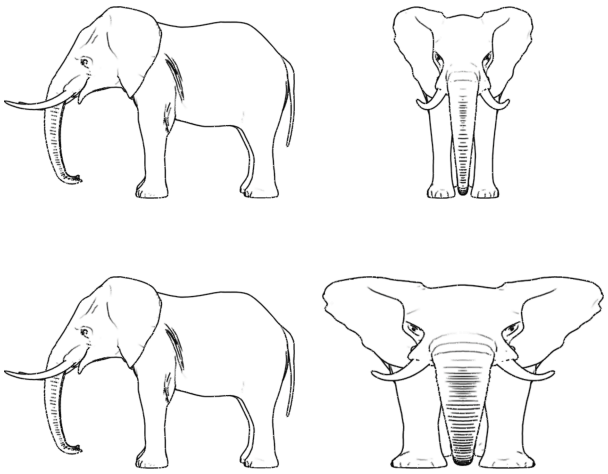
$$\# \max_{j=1,2,3} \pi_j(S_N) \geq N^{\frac{2}{3}}.$$

Exercise 2.3. 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 the results of this chapter 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 an elephant is overweight, there must be a way to place a mirror to make this obvious... as the name of this chapter suggests.

1. Notes, remarks and difficult questions

The inequality proved in this chapter, and the one in the next chapter, is a special case of the Loomis-Whitney inequality. The reader is encouraged to take a look at the original paper ([**15**]), a thoroughly readable article where this generalization is proved.

The result we considered suggests the following more general question. We have seen that at least one of the three coordinate planar projections of a finite point set must be reasonably large. What about a typical projection? Consider a vector $v \in \mathbb{R}^3$ of length one, and let H_v denote the plane through the origin perpendicular to v. Let $\pi_v(S)$ denote the projection of the set S onto H_v. Is there a sense in which $\#\pi_v(S)$ is "typically" large? Formulate a reasonable theorem and prove it!

Chapter 3

Projections in four dimensions

In this chapter we shall obtain an analog of Theorem 2.3 in four dimensions. As before, no background beyond high school mathematics is required and [**15**] is again recommended as the starting point for further reading. Things will get a little hairy however and more patience will be required to muscle through the details.

Unlike the previous chapter, it is difficult to motivate the four-dimensional problem by spraying elephants. However, our motivation need not involve physical dimensions. The idea is the same as before.

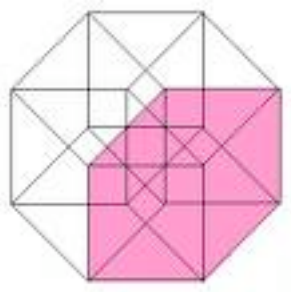


Figure 1. A depiction of a four dimensional cube

If the set is large, there must be a way to see it by taking lower dimensional "snapshots". This is where we are now headed.

Let $x = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4$. Define

$$\pi_{12}(x) = (x_1, x_2),$$

and define π_{ij}, $i \neq j$, $1 \leq i, j \leq 4$, in the same way. Let S_N be a subset of $\mathbb{R}^4$ containing N elements. Once again we ask, what is the relationship between the size of S_N and the sizes of $\pi_{ij}(S_N)$? In analogy with (2.2) we might expect an estimate of the form

$$\#S_N \leq (\Pi_{i<j;1\leq i,j\leq 4} \#\pi_{ij}(S_N))^{\alpha}, \tag{3.1}$$

for some $\alpha > 0$, where

$$\Pi_{1\leq i<j\leq 4} \#\pi_{ij}(S_N)$$

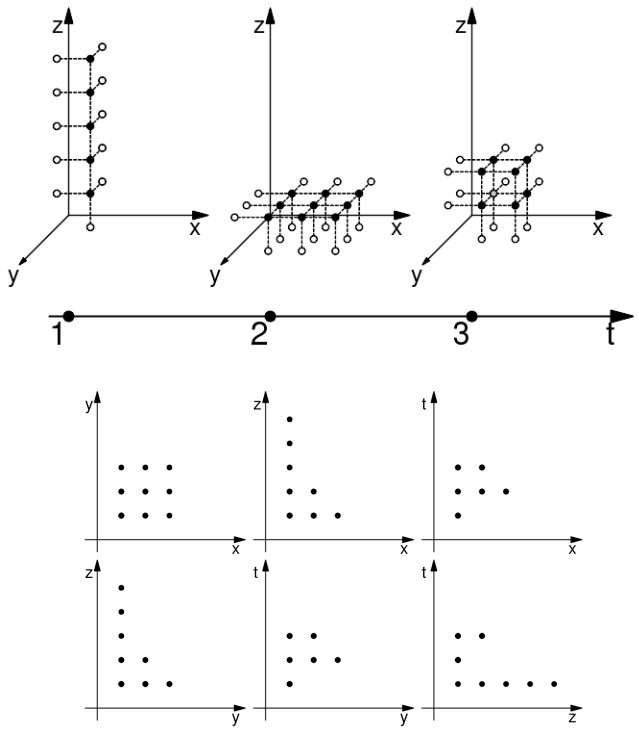
$$= \#\pi_{12}(S_N) \cdot \#\pi_{13}(S_N) \cdot \#\pi_{14}(S_N) \cdot \#\pi_{23}(S_N) \cdot \#\pi_{24}(S_N) \cdot \#\pi_{34}(S_N).$$

To be clear, notation $\Pi_{i\neq j;1\leq i,j\leq 4}$ means that we are taking a product over all the pairs (i, j) where $i < j$ and $1 \leq i, j \leq 4$. How many such pairs are there? We can just count and see that the number of such pairs is six. Alternatively, we could develop more of a robust counting mechanism which will serve us well later in the book. There are a total of 16 pairs (i, j), but since $i \neq j$ we are left with only 12 pairs. Among these pairs, for every (i, j) we also have a pair (j, i) and we only want one of them since we are interested in pairs where the second coordinate is larger than the first. This indeed leaves us with six pairs.

We are now in position to guess the exponent α in (3.1) by using what physicists call dimensional analysis. Since we are in four dimensions, it is reasonable to think of the left hand side of (3.1) as approximating four-dimensional volume, which we will measure in inches raised to the fourth power. The right hand side is the product of six terms, each of which is two dimensions, and thus measured in inches raised to the second power. We are thus led to the equation

$$inches^4 = \left((inches^2)^6\right)^{\alpha},$$

which forces us to conclude that $\alpha = \frac{1}{3}$. Indeed, we shall prove the following result.



Theorem 3.2. *With the notation above,*

$$\#S_N \leq (\Pi_{i<j;1\leq i,j\leq 4}\#\pi_{ij}(S_N))^{\frac{1}{3}}. \tag{3.3}$$

Exercise 3.1. Construct a set S_N for which the right hand side of (3.3) is much bigger than the left hand side. Similarly, construct a set for which the two sides are about equal. How does this compare with the examples and diagrams of the previous chapter?

To prove Theorem 3.2 we follow the blueprint of the previous chapter. We first observe that

$$\chi_{S_N}(x) \leq \Pi_{i<j;1\leq i,j\leq 4}\chi_{ij}(x_i, x_j), \tag{3.4}$$

which is true for the same reasons as before, namely that if $x \in S_N$, $\pi_{ij}(s) \in \pi_{ij}(S_N)$. This statement is so vacuous it almost hurts, but it does prove the assertion.

Exercise 3.2. Give an explicit example of a set S_N such that the estimate (3.4) is not sharp, namely that the right hand side can be equal to one for values of x for which the left hand side is 0. Do not stop there. For which sets is this inequality always sharp?

Let χ_{ij} denote the characteristic function of $\pi_{ij}(S_N)$. Recall that this means that χ_{ij} equals one on $\pi_{ij}(S_N)$ and 0 otherwise. We now apply (1.7) with $p = 3$ and $p' = \frac{3}{2}$ to the right hand side of (3.3) to see that

$$\#S_N \leq \sum_{x_1,x_2,x_3,x_4} \chi_{12}(x_1,x_2)$$

$$\cdot\chi_{13}(x_1,x_3)\chi_{14}(x_1,x_4)\chi_{23}(x_2,x_3)\chi_{24}(x_2,x_4)\chi_{34}(x_3,x_4)$$

$$\leq \left(\sum_{x_3,x_4} \chi_{34}^3(x_3,x_4)\right)^{\frac{1}{3}}$$

$$\cdot\left(\sum_{x_3,x_4}\left(\sum_{x_1,x_2} \chi_{12}(x_1,x_2)\dots\chi_{24}(x_2,x_4)\right)^{\frac{3}{2}}\right)^{\frac{2}{3}} = I \cdot II. \tag{3.5}$$

We have

$$I = (\#\pi_{34}(S_N))^{\frac{1}{3}}, \tag{3.6}$$

so we are down to estimating II. Almost there! Well, actually this is the hardest part of the proof. It must be because we have not done anything too interesting yet.

To estimate II we use a technique that is extremely important in modern harmonic analysis. We take advantage of the fact that while the exponent $\frac{3}{2}$ in the definition of II is very difficult to work with, exponents 1 and 2 are considerably easier. Our plan is to replace $\frac{3}{2}$ by 1 and 2, respectively, estimate those, and then come up with a scheme to recover the corresponding estimate with $\frac{3}{2}$. This technique falls under the general heading of "interpolation".

Well, enough talking... Let's get back to work. Let

$$F_s = \left(\sum_{x_3,x_4} \left(\sum_{x_1,x_2} \chi_{12}(x_1,x_2) \dots \chi_{24}(x_2,x_4) \right)^s \right)^{\frac{1}{s}}.$$

Consider the case $s = 1$. We have F_1 equals

$$\sum_{x_1,x_2,x_3,x_4} \chi_{12}(x_1,x_2)\chi_{13}(x_1,x_3)\chi_{14}(x_1,x_4)\chi_{23}(x_2,x_3)\chi_{24}(x_2,x_4).$$

Our goal is not to work too hard. We are already doing enough of that. The problem is that variables are all mixed together like an Irish stew and, to quote Ren and Stimpy, "No Sir, he (Mr. Horse, Alex Iosevich, whatever...) doesn't like it!". We resolve this difficulty by eliminating all the factors we do not like. By using the fact that $\chi_{ij}(x_i,x_j) \leq 1$, trivially, we see that

$$\begin{aligned} F_1 &\leq \sum_{x_1.x_2.x_3,x_4} \chi_{14}(x_1,x_4)\chi_{23}(x_2,x_3) \\ &= \sum_{x_1,x_4} \chi_{14}(x_1,x_4) \cdot \sum_{x_2,x_3} \chi_{23}(x_2,x_3) \\ &= \#\pi_{14}(S_N) \cdot \#\pi_{23}(S_N). \end{aligned} \tag{3.7}$$

To estimate F_2, we write F_2^2 in the form

$$\begin{aligned} &\sum_{x_3,x_4} \left(\sum_{x_1,x_2} \chi_{12}(x_1,x_2)\chi_{13}(x_1,x_3)\chi_{14}(x_1,x_4)\chi_{23}(x_2,x_3)\chi_{24}(x_2,x_4) \right)^2 \\ &= \sum_{x_1,x_2,x_1',x_2',x_3,x_4} A \cdot B \end{aligned} \tag{3.8}$$

where

$$A = \chi_{12}(x_1,x_2)\chi_{13}(x_1,x_3)\chi_{14}(x_1,x_4)\chi_{23}(x_2,x_3)\chi_{24}(x_2,x_4)$$

and

$$B = \chi_{12}(x_1',x_2')\chi_{13}(x_1',x_3)\chi_{14}(x_1',x_4)\chi_{23}(x_2',x_3)\chi_{24}(x_2',x_4).$$

We are again faced with the Irish stew problem and we resolve it in more or less the same way. By again using the fact that

$$\chi_{ij}(x_i,x_j) \leq 1,$$

we see by applying (3.8) that

$$\begin{aligned} F_2^2 &\leq \sum_{x_1,x_2,x_1',x_2',x_3,x_4} \chi_{12}(x_1,x_2)\chi_{13}(x_1',x_3)\chi_{24}(x_2',x_4) \\ &= \sum_{x_1,x_2} \chi_{12}(x_1,x_2) \cdot \sum_{x_1',x_3} \chi_{13}(x_1',x_3) \cdot \sum_{x_2',x_4} \chi_{24}(x_2',x_4) \\ &= \#\pi_{12}(S_N) \cdot \#\pi_{13}(S_N) \cdot \#\pi_{24}(S_N), \end{aligned}$$

from which it follows that

$$F_2 \leq \sqrt{\#\pi_{12}(S_N) \cdot \#\pi_{13}(S_N) \cdot \#\pi_{24}(S_N)}. \tag{3.9}$$

Where are we? We need to somehow go from having an upper bound (3.7) on F_1 and (3.9) on F_2 to a good upper bound on $F_{\frac{3}{2}}$.

3.1. Interpolation of estimates. What are we after, really... Suppose that $a_j \geq 0$,

$$\sum_j a_j \leq C_1, \tag{3.10}$$

and

$$\left(\sum_j a_j^2 \right)^{\frac{1}{2}} \leq C_2. \tag{3.11}$$

What can we say about

$$\left(\sum_j a_j^s \right)^{\frac{1}{s}} ? \tag{3.12}$$

It is reasonable for us to realize at this point that the estimate (1.8) may be useful because general exponents are involved. This means that we probably want to write

$$a_j^s = a_j^{s-\alpha} \cdot a_j^\alpha \tag{3.13}$$

and then apply (1.8) with some exponent p. The question is, what is α and what is p? Well, if we apply (1.8) to the right hand side of (3.13) we get

$$a_j^s = a_j^{s-\alpha} \cdot a_j^\alpha \leq \frac{a_j^{p(s-\alpha)}}{p} + \frac{a_j^{p'\alpha}}{p'}, \tag{3.14}$$

where $\frac{1}{p} + \frac{1}{p'} = 1$. Presumably we want

$$p(s - \alpha) = 1$$

and

$$p'\alpha = 2$$

since we wish to take advantage of (3.10) and (3.11).

Solving these equations we obtain

$$\alpha = 2(s - 1)$$

and

$$p = \frac{1}{2 - s}.$$

Plugging all this into (3.14), we obtain

$$a_j^s = a_j^{s-\alpha} \cdot a_j^\alpha$$

$$= a_j^{\frac{1}{p}} \cdot a_j^{\frac{2}{p'}} \leq \frac{a_j}{p} + \frac{a_j^2}{p'},$$

and, equivalently,

$$\frac{a_j^{\frac{1}{p}}}{C_1^{\frac{1}{p}}} \cdot \frac{a_j^{\frac{2}{p'}}}{C_2^{\frac{2}{p'}}} \leq \frac{a_j}{pC_1} + \frac{a_j^2}{p'C_2^2}. \tag{3.15}$$

Summing both sides in j we see that

$$\sum_j \frac{a_j^s}{C_1^{\frac{1}{p}} \cdot C_2^{\frac{2}{p'}}} = \sum_j \frac{a_j^{\frac{1}{p}} \cdot a_j^{\frac{2}{p'}}}{C_1^{\frac{1}{p}} \cdot C_2^{\frac{2}{p'}}}$$

$$\leq \frac{1}{p} \frac{\sum_j a_j}{C_1} + \frac{1}{p'} \frac{\sum_j a_j^2}{C_2^2} = \frac{1}{p} + \frac{1}{p'} = 1.$$

Multiplying both sides by $C_1^{\frac{1}{p}} \cdot C_2^{\frac{2}{p'}}$ and plugging in $p = \frac{1}{2-s}$, we obtain the following attractive looking result.

Theorem 3.16. *Let a_j be a sequence of positive numbers. Let $1 < s < 2$. Suppose that (3.10) and (3.11) hold. Then*

$$\left(\sum_j a_j^s \right)^{\frac{1}{s}} \leq C_1^{\frac{2}{s}-1} \cdot C_2^{2-\frac{2}{s}}.$$

The proof above looks innocent enough, but let us do a bit of a postmortem. Why did we perform the division "trick" in (3.15)? This convenient technique is called scaling. We looked ahead and saw that $\sum_j a_j = C_1$ and $\sum_j a_j^2 = C_2^2$, by assumption, and divided by the appropriate factors accordingly.

Corollary 3.17. *Under the assumptions of Theorem 3.16 we have*

$$\left(\sum_j a_j^{\frac{3}{2}} \right)^{\frac{2}{3}} \leq C_1^{\frac{1}{3}} \cdot C_2^{\frac{2}{3}}.$$

We are now ready to tackle the projection problem again. Using Corollary 3.17 we see that

$$F_s \leq (\#\pi_{12}(S_N) \cdot \#\pi_{13}(S_N) \cdot \#\pi_{14}(S_N) \cdot \#\pi_{23}(S_N) \cdot \#\pi_{24})^{\frac{1}{3}}.$$

Combining this with (3.5) and (3.6) we see that $\#S_N \leq$

$$(\#\pi_{12}(S_N) \cdot \#\pi_{13}(S_N) \cdot \#\pi_{14}(S_N) \cdot \#\pi_{23}(S_N) \cdot \#\pi_{24}(S_N) \cdot \#\pi_{34}(S_N))^{\frac{1}{3}},$$

thus verifying Theorem 3.2.

Exercise 3.3. Another way of approaching the main result of this section is the following. First, prove a corresponding result for three-dimensional projections. Then apply the results of the previous section to extract two-dimensional projections from the three-dimensional ones. Carry out the details of this calculation and make sure that the resulting theorem agrees with the one we proved in this chapter.

1. Notes, remarks, and difficult questions

Well... we have now analyzed two-dimensional projections in three dimensions and two-dimensional projections in four dimensions. What about k-dimensional projections in d dimensions? The aforementioned Loomis-Whitney inequality deals precisely with this general situation.

One way to guess what the general (k, d) projection theorem should say is to use dimensional analysis as we did above. The left hand side of the inequality should probably be the size of S. In d dimensions we can measure it in inches raised to the power of d. On the

right hand side of the inequality we should presumably have a product of all possible k-dimensional coordinate projections in $\mathbb{R}^d$ raised to some power. How do we determine this power? Well, each k-dimensional projection can be measured in inches raised to the power of k. How many different k-dimensional coordinate projections are there? As many as there are ways of choosing k objects out of d. This number, which equals $\frac{d!}{k!(d-k)!}$, is discussed extensively in Chapter 9 later in this book. Now find the power to which you must raise the right hand side in order to make the units match!

This reasoning will give you the statement of the general (k, d) projection theorem. Can you prove it? At the very least, can you figure out which version of the Hölder inequality is involved in the (k, d) case?

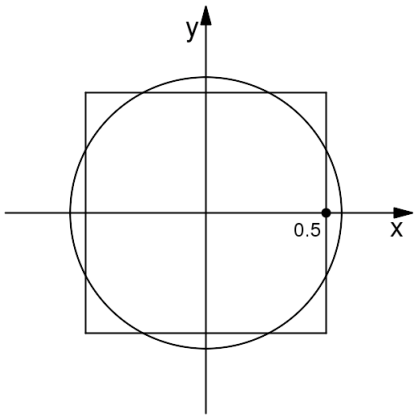
Chapter 4

Projections and Cubes

In this chapter we build on the intuition of the previous two chapters and explore yet another aspect of projections. Knowledge of higher dimensional calculus is helpful here, though if you are only familiar with double and triple integrals, you will be fine if you try hard! Given a subset of $\mathbb{R}^d$ we attempt to determine the largest possible projection of this set that contains the cube of side-length one. Knowledge of first and second year calculus is required for this chapter.

The basic idea of what we are about to do is the following. We take a convex subset of d-dimensional Euclidean space $\mathbb{R}^d$ and try to fit the cube of side-length one inside. If this is not possible, we

Figure 1. Balls inside cubes are common, but the other way around ?!



try to fit the cube of side-length one inside of one of its coordinate projections. For example, consider the disc of area one in the plane. Its radius is $\frac{1}{\sqrt{\pi}}$. It is not difficult to see that the cube of side-length one does not fit inside the disc. However, the projection of this disc onto the x-axis is the interval of length $\frac{2}{\sqrt{\pi}} > 1$, so the unit "square" in one dimension fits inside of it really well!

This simple calculation raises an amusing question of what happens if $S = B_d$, the d-dimensional ball of "volume" one... There are several questions to be answered here, not the least of which is what is the "volume" of a d-dimensional ball, and, if this volume is one, what is the radius! Let's start exploring, which means that we must get more precise.

Let $S \subset \mathbb{R}^d$. Let $x = (x_1, \ldots, x_d)$. Let J be a subset of $\{1, 2, \ldots, d\}$ and let $P_J(x) = (x_{i_1}, x_{i_2}, \ldots, x_{i_k})$, where $\{i_1, \ldots, i_k\} = J$. We are going to define a peculiar notion of "dimension" of S as the largest positive integer k such that $P_J(S)$ contains a cube of side-length 1 and $\#J = k$. We shall not dwell on applications of this notion here, but we shall make some calculations that are both fascinating and instructive.

First, what is the "volume" of the ball of radius R in $\mathbb{R}^d$? Presumably it is

$$\int_{|x|\le R} dx = \int_0^R \int_{S^{d-1}} d\omega \; r^{d-1} dr = R^d \frac{|S^{d-1}|}{d}, \tag{4.1}$$

where $d\omega$ is the "surface" measure on

$$S^{d-1} = \left\{ x \in \mathbb{R}^d : |x| = \sqrt{x_1^2 + \cdots + x_d^2} = 1 \right\},$$

and $|S^{d-1}|$ is the "area" of S^{d-1}.

Exercise 4.1. If you are only familiar with two-dimensional and three-dimensional calculus, try to use those methods to verify (4.1). In other words, explain how "polar" coordinates are being used here.

4.1. Radius calculation. It follows that if we are interested in a ball of "volume" 1, we must solve the equation

$$(R_d)^d \frac{|S^{d-1}|}{d} = 1,$$

which yields

$$R_d = \left(\frac{d}{|S^{d-1}|} \right)^{\frac{1}{d}}, \tag{4.2}$$

which only shifts the problem for the moment because we have no idea what $|S^{d-1}|$ is, so let's get to it. We have

$$I_d = \int_{\mathbb{R}^d} e^{-|x|^2} dx = \int_{-\infty}^{\infty} e^{-x_1^2} dx_1 \cdot \cdots \cdot \int_{-\infty}^{\infty} e^{-x_d^2} dx_d$$

$$= \left(\int_{-\infty}^{\infty} e^{-t^2} dt \right)^d = I_1^d.$$

On the other hand, if we employ "polar" coordinates once again, define

$$I_d = \int_0^{\infty} \int_{S^{d-1}} d\omega e^{-r^2} r^{d-1} dr$$

$$= |S^{d-1}| \cdot \int_0^{\infty} e^{-r^2} r^{d-1} dr = |S^{d-1}| \cdot G_d,$$

with G_d being defined by this relation.

We conclude that

$$|S^{d-1}| = \frac{I_1^d}{G_d},$$

which would be worth something if we compute both I_1 and G_d. Fortunately, we can.

In order to compute I_1 we use polar coordinates again, but only in two dimensions. We have

$$\begin{aligned} I_1^2 &= \int_{-\infty}^{\infty} e^{-x^2} dx \cdot \int_{-\infty}^{\infty} e^{-y^2} dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy \\ = \int_0^{\infty} \int_0^{2\pi} d\theta e^{-r^2} r dr &= 2\pi \cdot \int_0^{\infty} e^{-r^2} r dr = \pi. \end{aligned}$$

It follows that

$$I_1 = \sqrt{\pi}. \tag{4.3}$$

Now, G_d is a tougher nut to crack, but we should be able to do something. Observe that we can make a change of variables $u = r^2$ to see that

$$G_d = \frac{1}{2} \int_0^{\infty} e^{-u} u^{\frac{d}{2}-1} du.$$

Integrating by parts once, we see that

$$\begin{aligned} G_d &= \frac{1}{2} \cdot \left(\frac{d}{2} - 1\right) \int_0^{\infty} e^{-u} u^{\frac{d-2}{2}-1} du \\ &= \left(\frac{d}{2} - 1\right) G_{d-2}. \end{aligned} \tag{4.4}$$

This does not look incredibly encouraging, but we will make it better by assuming that d is even. Let $d = 2n$ and define

$$H_n = G_{2n}.$$

It follows from (10.12) that

$$H_n = (n-1) H_{n-1},$$

and this we can work with. We have

$$\frac{H_k}{H_{k-1}} = k - 1,$$

so

$$\frac{H_2}{H_1} \cdot \frac{H_3}{H_2} \cdot \cdots \cdot \frac{H_n}{H_{n-1}} = 1 \cdot 2 \cdot \cdots \cdot (n-1) = (n-1)!,$$

so

$$\frac{H_n}{H_1} = (n-1)!,$$

and we are down to figuring out what H_1 is. This we are going to compute directly. We have

$$H_1 = G_2 = \frac{1}{2} \int_0^\infty e^{-u} du = \frac{1}{2}.$$

It follows that

$$H_n = \frac{1}{2}(n-1)!.$$

And, finally, we have that if $d = 2n$,

$$|S^{d-1}| = \frac{2\pi^n}{(n-1)!},$$

so

$$\frac{|S^{d-1}|}{d} = \frac{2\pi^n}{2n(n-1)!} = \frac{\pi^n}{n!}.$$

Plugging this back into (4.2) we see that

$$R_d = \left(\frac{n!}{\pi^n}\right)^{\frac{1}{2n}} = \frac{(n!)^{\frac{1}{2n}}}{\sqrt{\pi}}. \tag{4.5}$$

Exercise 4.2. Another way to compute the volume of the ball of radius one is the following. Let $\omega_d(R)$ denote the volume of the ball of radius R in $\mathbb{R}^d$. Check that

$$\omega_d(R) = R^d \omega_d(1).$$

Next observe that

$$\omega_d(1) = \int_{x_1^2+\cdots+x_d^2 \leq 1} dx = \int_0^1 \int_{x_1^2+\cdots+x_{d-1}^2+t^2 \leq 1} dx' dt,$$

where

$$dx' = dx_1 dx_2 \ldots dx_{d-1}.$$

It follows that

$$\omega_d(1) = \int_0^1 \int_{x_1^2+\cdots+x_{d-1}^2 \leq 1-t^2} dx' dt$$

$$= \int_0^1 \omega_{d-1}(\sqrt{1-t^2})dt = \omega_{d-1}(1)\int_0^1 (1-t^2)^{\frac{d-1}{2}}dt.$$

Integrate by parts repeatedly and complete the calculation.

4.2. Back to projections. Now that we know the radius of the ball of volume one, we are ready for the exciting task of fitting a cube of side-length one into its projections. A beautiful thing about a ball is that all of its projections are balls. If $\#J = k$, the ball is k-dimensional and is of radius R_d. Let s be the side-length of the cube centered at the center of the k-dimensional ball of radius R_d. By the Pythagorean theorem

$$k\Big(\frac{s}{2}\Big)^2 = R_d^2,$$

so

$$s = \frac{2R_d}{\sqrt{k}}.$$

Since we must have $s \geq 1$, we conclude that

$$k \leq 4R_d^2 = \frac{4}{\pi}(n!)^{\frac{1}{n}}. \tag{4.6}$$

Exercise 4.3. How does the estimate in (4.6) change if we replace a cube of side-length one in our procedure by a cube of side-length ϵ?

In a way, we have the estimate we sought. Unfortunately, it is not incredibly clear what $(n!)^{\frac{1}{n}}$ and this is what we are after now.

4.3. Estimating factorials. We are not going to be incredibly formal here, just precise enough to get the job done. First, we need a definition. We say that

$$a(T) \approx b(T),$$

with respect to a variable T if

$$\frac{a}{b} \to 1 \text{ as } T \to \infty.$$

With this in mind, observe that

$$\log(n!) = \sum_{k=1}^{n} \log(k).$$

Suppose that we could show that

$$\sum_{k=1}^{n} \log(k) \approx \int_1^n \log(t)dt. \tag{4.7}$$

We could then conclude that

$$\log(n!) \approx n\log(n) - n,$$

from which it follows that

$$n! \approx \left(\frac{n}{e}\right)^n. \tag{4.8}$$

This would indeed be very nice because we could then insert this into (4.6) and conclude that

$$k \leq \frac{4}{\pi}(n!)^{\frac{1}{n}} \approx \frac{4}{\pi} \cdot \frac{n}{e} = \frac{2d}{\pi e},$$

an attractive estimate indeed.

It remains to verify (4.7). On one hand,

$$\int_1^n \log(t)dt = \sum_{k=2}^{n} \int_{k-1}^{k} \log(t)dt \tag{4.9}$$

$$\leq \sum_{k=2}^{n} \log(k) = \sum_{k=1}^{n} \log(k),$$

since $\log(t)$ is an increasing function and $\log(1) = 0$.

On the other hand, the right hand side of (4.9) is bounded from below by

$$\sum_{k=2}^{n} \log(k-1) = \sum_{k=1}^{n-1} \log(k)$$

$$= \sum_{k=1}^{n} \log(k) - \log(n).$$

It follows that

$$1 - \frac{\log(n)}{\sum_{k=1}^{n} \log(k)} \leq \frac{\int_1^n \log(t)dt}{\sum_{k=1}^{n} \log(k)} \leq 1,$$

and (4.7) follows if we can show that

$$\frac{\log(n)}{\sum_{k=1}^{n} \log(k)} \to 0 \text{ as } n \to \infty. \tag{4.10}$$

Since

$$\sum_{k=1}^{n} \log(k) = \sum_{k=2}^{n} \log(k) \geq \log(2)(n-1),$$

it follows that

$$0 \leq \frac{\log(n)}{\sum_{k=1}^{n} \log(k)} \leq \frac{\log(n)}{\log(2)(n-1)},$$

so it suffices to show that

$$\frac{\log(n)}{n-1} \to 0 \text{ as } n \to \infty.$$

There are many ways to see this, using the mean value theorem from calculus, for example, but we prefer to take a slightly different route. Observe that

$$\log(n) = \int_1^n \frac{dt}{t} = \int_1^R \frac{dt}{t} + \int_R^n \frac{dt}{t} = I + II.$$

Now,

$$I \leq (R-1)$$

because $\frac{1}{t}$ is never larger than 1 on the interval $[1, R]$. By the same reasoning,

$$II \leq (n-R) \cdot \frac{1}{R}.$$

The two expressions are equal if we set $R = \sqrt{n}$. It follows that

$$\log(n) \leq 2(\sqrt{n}-1),$$

and

$$\frac{\log(n)}{n-1} \leq \frac{2(\sqrt{n}-1)}{n-1} \to 0 \text{ as } n \to \infty.$$

We conclude that (4.10) and thus (4.7) are verified! Remember what all of this was for. In (4.6) we computed the "dimensions" of the unit ball in terms of an expression involving factorials and roots and the purpose of the rest of the chapter was to get a tighter grip on what exactly that expression says. In the process we discovered that, as $d \to \infty$, the "dimension" of the unit ball is $\approx d \cdot \frac{2}{\pi e}$, which is close to $\frac{d}{4}$. In other words, we do not need to project the ball very far down in order to be able to fit the unit cube inside. This is not so obvious, is it?

Exercise 4.4. Carry out the details of the following proof of the fact that $\lim_{n\to\infty} \frac{\log(n)}{n} = 0$. First prove that the sequence $\frac{\log(n)}{n}$ is decreasing (by taking a derivative, for example). Conclude that it suffices to establish the limit for a subsequence. In particular, it is enough to take $n = 2^k$, which reduces the limit we are after to showing that

$$\lim_{k\to\infty} \frac{k}{2^k} = 0. \tag{4.11}$$

Exercise 4.5. The argument above shows that if $\alpha > \frac{1}{2}$ then

$$\frac{\log n}{n^\alpha} \to 0 \text{ as } n \to \infty.$$

However, the mean value theorem (or L'Hopital's rule, as some people insist on calling it) easily shows that this should hold with any $\alpha > 0$. Can you refine the procedure used above to show this? Yes you can... so just do it! In an effort to be slightly more helpful, we decomposed the integral into two pieces to prove that $\log(n) \leq 2(\sqrt{n}-1)$. How many pieces do you think it is going to take to obtain a similar inequality with $n^{\frac{1}{3}}$ on the right hand side?

Let S be a finite set containing k elements. Prove that the total number of subsets of S is 2^k. To see this observe that you can walk up to each element of S and ask: "ARE YOU IN A GIVEN SUBSET, YES OR NO?!" The answer is YES or NO for each of the k elements, which results in 2^k possible choices, as claimed.

Next observe that the number of subsets of S of size 2 is $\frac{k(k-1)}{2}$. To see this, note we have k choices for the first of two elements of such a subset, and only $k-1$ choices for the second because the elements must be different. We divide by two because the set does not become a different set if the order of the two elements is switched. We conclude that the number of subsets of S of size 2 is indeed $\frac{k(k-1)}{2}$ from which we deduce that

$$\frac{k(k-1)}{2} < 2^k,$$

since, after all, there are fewer subsets of size 2 than the total number of subsets... Plugging this deep observation into (4.11) we see that

$$\frac{k}{2^k} \leq \frac{k}{\frac{k(k-1)}{2}} = \frac{2}{k-1} \to 0 \text{ as } k \to \infty.$$

Exercise 4.6. Can you improve upon (4.8) and show that in fact

$$n! \approx \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n ?$$

Hint: Show that

$$n! = \int_0^\infty e^{-t} t^n dt.$$

Observe that the critical point (the point where the derivative changes sign) of the function $e^{-t}t^n$ is at $t = n$... and go from there. This is not entirely trivial and is known as Stirling approximation. See, for example, [**7**].

Exercise 4.7. Let $1 < p < \infty$ and define

$$B_{d,p}(r) = \{x \in \mathbb{R}^d : |x_1|^p + \cdots + |x_d|^p \leq r^p\}.$$

Let $r = r_d$ such that the volume of $B_{d,p}(r_d)$ equals 1. What is the dimension of this set in the sense of this chapter?

Suffering is unavoidable here... so please do not start complaining if you are not done in two or three hours... The hard part is not the computation of the volume of $B_{d,p}(r)$, which is essentially the same as the corresponding procedure for $B_{d,2}(r)$. Fitting a cube inside $B_{d,p}(r)$ is a different matter. This part will require some pain.

1. Notes, remarks and difficult questions

So what have we seen in this chapter? We must project the ball of volume one in $\mathbb{R}^d$ onto an approximately $\frac{d}{4}$ coordinate sub-space before it contains the unit cube. Can you construct a family of sets $K_d \subset \mathbb{R}^d$ of volume one such that the "dimension", in the sense of this chapter, is substantially smaller? Can you construct these sets such that the dimension is $\approx \log(d)$? If you cannot, can you prove that this is never the case? If you can, is it possible to put some reasonable geometric restrictions on K_d such that the dimension is $\geq cd$?

Another point worth considering is the fact that we only really analyzed the case when d is even. Can you go through the chapter and work out the case when d is odd?

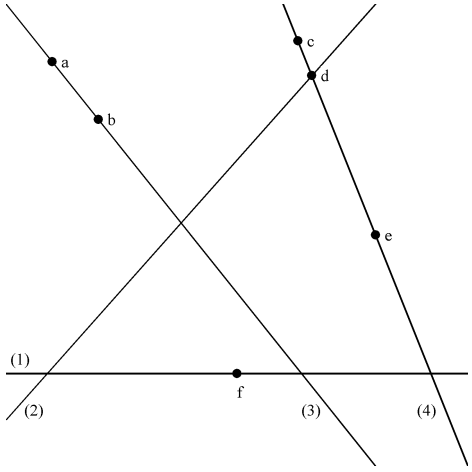
Does the problem become substantially different if instead of coordinate projections we consider sections? More precisely, suppose that we ask for the dimension of the largest coordinate hyper-plane H such that $K_d \cap H$ contains the unit cube? In the case of the ball, sections and projections correspond, but in general this is not the case. Can you construct an example of a family of convex sets of $K_d \subset \mathbb{R}^d$ of volume one such that the section notion of dimension leads to vastly different results than the projection notion of dimension? If you can, what is going on? If you cannot, can you formulate and prove a general theorem?

Chapter 5

Incidences and matrices

This chapter requires no background at all beyond high school mathematics. The reader is encouraged to look at the original papers by Szemeredi and Trotter ([**18**]), and by Bourgain, Katz and Tao ([**3**]) where the best known incidence theorems are proved in various contexts. The goal of this chapter is to present the idea of incidences which arise almost inevitably in most if not all areas of mathematics. The basic idea is the following. Let P be a finite set of points and let L be a finite set of geometric objects. We are being ridiculously vague, but this is done for a reason that will become apparent later. A pair (p, l) is called an incidence if $p \in P$, $l \in L$, and p lies on l. What is the total number of incidences? Since we are counting pairs it is clear that this number cannot exceed $\#P \cdot \#L$, and if the story ended there, you would be justifiably disappointed. However, it turns out that under a variety of reasonable assumptions on points, geometric objects and the structure of the space where these creatures reside, the upper bound on the number of incidences is much more restricted and beautiful consequences follow.

It is time to get serious once more. Consider a set of n lines and n points in the plane. As before, define an incidence to be a pair (p, l), where p is one of the points in our point set, l is one of the lines in our set of lines, and p lies on l. Let $I(n)$ denote the total number of incidences determined by a given set of n points and a given set of



n lines. In order to avoid needless headaches we assume that every point in our point set lies on at least one line in our set of lines, and every line in our line set contains at least one point in our point set.

How large can $I(n)$ be? Well, it is clear that $I(n) \leq n^2$. This observation is not terribly valuable, however, since $I(n)$ cannot possibly be this large! I mean, how can every line contain every point, and every point lie on every line?! You might retort that maybe, just maybe, it is possible for about $n/10$ lines to contain about $n/100$ points each, and for each of those points to be contained in about $n/1000$ of those lines. We shall see that nothing like that can happen.

Our main tools in this endeavor are matrices and the C-S inequality. Recall that an n by n matrix A is an array with n rows and n columns. The elements of A are designated by a_{ij}, where i determines the row and j determines the column. Let's define A as follows. Enumerate the n points in our point set from 1 to n, and do the same for lines in our set of lines. Let $a_{ij} = 1$ if the i'th point lies on the j'th line, and 0 otherwise. Observe that if j and j' are fixed, with $j \neq j'$,

$$a_{ij} \cdot a_{ij'} = 1 \tag{5.1}$$

for at most one value of i. This is because $a_{ij} \cdot a_{ij'} = 1$ if and only if $a_{ij} = 1$ and $a_{ij'} = 1$. This means that the i'th point is on the

j'th line and also on the j'th line. Intersection of two distinct lines is either empty or consists of exactly one point. It follows that indeed the equality in (5.1) can hold for at most one i.

We are now ready for action. What is $I(n)$? It is nothing more than the total number of 1s in A! Since A consists of only 1s and 0s,

$$I(n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}\right)\cdot 1$$

$$\leq \left(\sum_{i=1}^{n} 1\right)^{\frac{1}{2}}\left(\sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}\right)^2\right)^{\frac{1}{2}} = \sqrt{n}\cdot\left(\sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}\right)^2\right)^{\frac{1}{2}},$$

where in the second line, we used the Cauchy-Schwarz inequality.

Now,

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}\right)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{j'=1}^{n} a_{ij}a_{ij'}$$

$$= \sum_{i=1}^{n}\sum_{\{j\neq j':1\leq j,j'\leq n\}} a_{ij}a_{ij'} + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2 = apple + orange.$$

Observe that the second sum corresponds to the case when $j = j'$, which is why the parameter j' does not appear there.

To estimate apple we use (5.1). Indeed, since $a_{ij}\cdot a_{ij'} = 1$ for at most one i,

$$apple \leq \#\{(j,j') : 1\leq j,j'\leq n; j\neq j'\} = n^2 - n.$$

On the other hand,

$$orange \leq \#\{(i,j) : 1\leq i,j\leq n\} = n^2.$$

Putting everything together and using the fact that $n^2 - n \leq n^2$, we see that

$$I(n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} \leq \sqrt{2}\cdot n^{\frac{3}{2}}.$$

We conclude that the number of incidences between n points and n lines in the plane is at most $\sqrt{2}n^{\frac{3}{2}}$. Can this estimate be improved? Sure it can... The sharp answer is $I(n) \leq Cn^{\frac{4}{3}}$, where C is a fixed

positive constant. This is the celebrated Szemeredi-Trotter incidence theorem ([**18**]) and it is sharp in the sense that one can construct a set of n lines and n points such that the number of incidences is approximately $n^{\frac{4}{3}}$, up to a constant. The proof of this result will appear later in this book.

Exercise 5.1. Read Chapter 6 and 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 as your point set all the points in $\mathbb{F}_q^2$ and take as your line set all the lines in $\mathbb{F}_q^2$.

Exercise 5.2. Let S_N be a subset of the plane consisting of N points. Define $\Delta(S_N)$ to be the set of pair-wise distances, namely

$$\left\{ \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} : x = (x_1, x_2) \in S_N, y = (y_1, y_2) \in S_N \right\}.$$

Use the results of this chapter to show that $\#\Delta(S_N) \geq C\sqrt{N}$ for some constant C independent of N. Can you do better? The conjectured answer is that

$$\#\Delta(S_N) \geq C \frac{N}{\sqrt{\log(N)}}.$$

The best known result to date, due to Katz and Tardos ([**13**]), based on the previous result due to Solymosi and Toth ([**19**]) is

$$\#\Delta(S_N) \geq CN^{\beta},$$

where $\beta \approx .86$ (See [**13**]).

What about higher dimensions? If $S_N \subset \mathbb{R}^d$ of size N, prove that

$$\#\Delta(S_N) \geq CN^{\frac{1}{d}}.$$

Can you do better? The conjectured answer here is

$$\#\Delta(S_N) \geq CN^{\frac{2}{d}}$$

in dimensions three and higher. Do you see where the exponent $\frac{2}{d}$ is coming from? Hint: Let

$$S_N = \{n = (n_1, \ldots, n_d) : n_j \in \mathbb{Z}; 1 \leq n_j \leq N^{\frac{1}{d}}\}.$$

Exercise 5.3. 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 5.4. Prove that n points and n spheres of the same radius in $\mathbb{R}^d$, $d \geq 4$, can have n^2 incidences. Use techniques of the chapter to see that when $d = 2$ the number of incidences is $\leq Cn^{\frac{3}{2}}$. What can you say about the case $d = 3$?

1. Notes, remarks and difficult questions

The best known incidence theorem in the plane is the Szemeredi-Trotter incidence theorem which says that the number of incidences between n points and m lines is at most

$$C(n + m + (nm)^{\frac{2}{3}}). \tag{5.2}$$

Use it to prove the following interesting fact. Let $A \subset \mathbb{Z}$ and define

$$A + A = \{a + a' : a, a' \in A\}$$

and

$$A \cdot A = \{a \cdot a' : a, a' \in A\}.$$

Then

$$\max\{\#(A + A), \#(A \cdot A)\} \geq C(\#A)^{\frac{5}{4}}.$$

Hint: For each $a, a' \in A$, consider a line $\{(at, t + a')\}$. As your set of points use $(A \cdot A) \times (A + A)$. Now apply (5.2) and go from there.

How does one prove the Szemeredi-Trotter incidence theorem itself? Try to muscle through the following outline. Rudimentary knowledge of graph theory would be helpful but not entirely necessary. You should be warned, however, that if you have not seen any

graph theory before, you will probably suffer. A graph is a finite set G equipped with a function

$$E : \{(x, y) \in G \times G : x \neq y\} \to \{0, 1\}.$$

You should think of $E(x, y) = 1$ as corresponding to the notion that x and y are connected by an edge, and $E(x, y) = 0$ as corresponding to the notion that x and y are not connected by an edge.

A *drawing* of a graph is a collection of $\#G$ points in the plane such that each pair of these points (x, y) is connected by a piece-wise differentiable curve (not necessarily a straight line segment) if and only if $E(x, y) = 1$. The representations of points of G in the drawing are called vertices. We say that the graph is *planar* if there exists a drawing of G such that edges do not intersect except at the vertices. We say that a *face* is a region in the drawing of a graph bounded by edges. Observe the following.

Exercise 5.5. Suppose that G is a planar graph with n vertices, e edges and f faces. Then

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

Think induction...

Exercise 5.6. Armed with the previous exercise, prove, under the same assumptions, that

$$3f \leq 2e$$

and conclude that if G is planar, then

$$e \leq 3n - 6.$$

Define a crossing, in a drawing of a graph, to be the intersection of edges not at a vertex. If G is not a planar graph, we define the *crossing number* of G, denoted by $cr(G)$, to be the smallest number of crossings over all the possible drawings of G.

Exercise 5.7. Use Exercise 5.6 to see that if G is not planar, then

$$cr(G) \geq e - 3n + 6, \tag{5.3}$$

where, as before, n is the number of vertices and e is the number of edges.

With all this technology in tow, we present our ballistic missile.

Lemma 5.4. *Suppose that $e \geq 4n$. Then*

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

Before we outline the proof of Lemma 5.4, let us see how it implies (5.2). Draw the family of lines and points in question. Make the points vertices of our draft. Connect two vertices by an edge if they are *consecutive* on some line.

Exercise 5.8. Prove that

$$e = I - m,$$

where m is the number of lines and I is the number of incidences.

Hint: Induction...

We are now thoroughly in business. Either $e \leq 4n$, in which case,

$$I \leq m + 4n,$$

which is even better than (5.2), or $e > 4n$, in which case Lemma 5.4 applies and we have

$$cr(G) \geq c\frac{e^3}{n^2} = c\frac{(I-m)^3}{n^2}. \tag{5.5}$$

To get any mileage out of this, though, we need an upper bound on $cr(G)$. Well, where do crossings come from? They come from edges intersecting. Where to edges come from? They come from lines. It follows that

$$cr(G) \leq m^2. \tag{5.6}$$

Combining (5.5) and (5.6) we see that

$$\frac{(I-m)^3}{n^2} \leq m^2,$$

which implies (5.2) instantly.

We are thus reduced to proving Lemma 5.4. Here is an idea. Consider a random subgraph of G where each vertex is chosen with probability p.

Exercise 5.9. The expected number of surviving vertices is np, the expected number of surviving edges is ep^2 and the expected value of the crossing number is $\leq p^4 cr(G)$.

Now plug these into (5.3), optimize in p, and finish the proof. This is it! We have outlined the proof of the Szemeredi-Trotter incidence theorem. Work through all the details, though! No loafing....

Chapter 6

Basics of grids over finite fields

Mathematicians sometimes construct simplified models in order to understand and illustrate the salient features of the problems we are studying. In the past two decades, vector spaces over finite fields, or grids over finite fields as we call them below, have been used to study analogs of important problems in the Euclidean space. Take a look at a beautiful survey by Tom Wolff ([**21**]) for a description of such problems. Why is this interesting or important? The main point is that in many instances grids over finite fields allow us to focus on the essence of the problem at hand without simultaneously dealing with the technical complications arising from the structure of the Euclidean space. Unfortunately, or fortunately, finite fields also bring their own arithmetic complications into the mix. More about that later.

Let q be a positive integer, prime in the sense that an integer a divides q if and only if $a = 1$ or $a = q$. Define $\mathbb{F}_q$ to be the set

$$\{0, 1, 2, \ldots, (q-1)\}$$

with the rule that addition and multiplication is taken modulo q. What this means is that if $a \in \mathbb{F}_q$ and $b \in \mathbb{F}_q$, $a+b$ (in the world of $\mathbb{F}_q$) is obtained by adding a and b in the standard way and computing the remainder after division by q. Similarly, to compute $a \cdot b$, we multiply

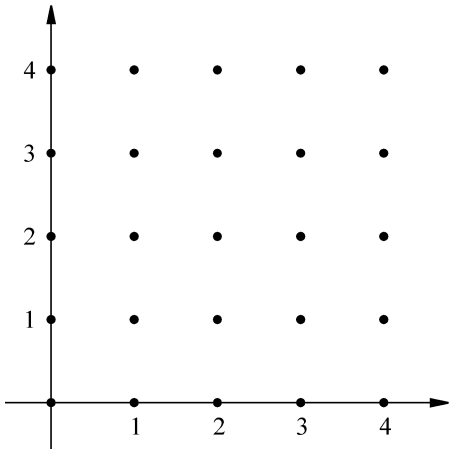


Figure 1. The grid $\mathbb{F}_5^2$.

a and b in the standard way and again compute the remainder after the division by q.

Example 6.1. Let $q = 5$. Then $\mathbb{F}_5 = \{0, 1, 2, 3, 4\}$. Suppose that we want to multiply 2 and 4 in the world of $\mathbb{F}_5$. Well, $2 \cdot 4 = 8$ in the sense of regular multiplication. If we divide 8 by 5, the remainder is 3. Thus $2 \cdot 4 = 3$ in the world of $\mathbb{F}_5$.

Now let's compute $2 + 4$. In the sense of regular addition, this equals 6. The remainder of division of 6 by 5 is 1. Thus $2 + 4 = 1$ in the world of $\mathbb{F}_5$.

We shall eventually stop saying "in the world of $\mathbb{F}_q$". We shall simply perform our addition and multiplication according to the rules we just described and illustrated. Relapses are possible, of course...

We now introduce a "grid" $\mathbb{F}_q^d$, defined as a set of d-tuples

$$(a_1, a_2, \ldots, a_d),$$

such that a_j is an element of $\mathbb{F}_q$.

Example 6.2. The set $\mathbb{F}_3^2$ consists of 9 pairs: $(0,0)$, $(0,1)$, $(0,2)$, $(1,0)$, $(1,1)$, $(1,2)$, $(2,0)$, $(2,1)$, and $(2,2)$.

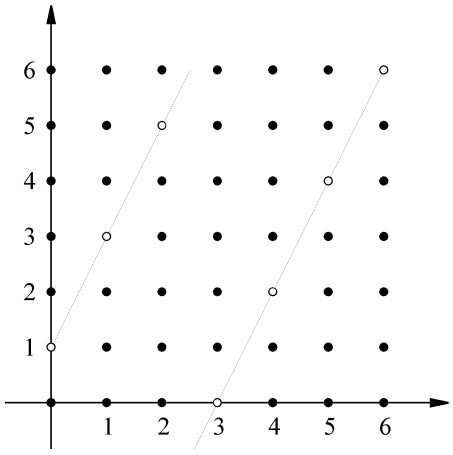


Figure 2. A line in $\mathbb{F}_7^2$.

The set $\mathbb{F}_3^3$ consists of 27 triples. Write them out!

Exercise 6.1. Show that $\mathbb{F}_q^d$, d a positive integer, consists of q^d elements.

We now introduce the notion of a line in $\mathbb{F}_q^d$. Let $x \in \mathbb{F}_q^d$, and let $v \in \mathbb{F}_q^d$, with the additional restriction that $v \neq (0, 0, \ldots, 0)$. Let

$$L(x, v) = \{x + tv : t = 0, 1, \ldots, q-1\}.$$

Example 6.3. Let $x = (0, 1)$ and $v = (1, 1)$ and let $q = 3$. Then $L(x, v)$ is the "line" consisting of points $(0, 1)$, $(1, 2)$, and $(2, 0)$. Why is that? Well, by definition, $L((0, 1), (1, 1)) = \{(0, 1) + t(1, 1) : t = 0, 1, 2\}$. If $t = 0$, we get $(0, 1)$ easily enough. If $t = 1$, we get $(1, 2)$ with no problem. If $t = 2$, we get $(0 + 2 \cdot 1, 1 + 2 \cdot 1) = (2, 0)$ since $1 + 2 \cdot 1 = 3$ which is 0 in the world of $\mathbb{F}_3$.

Exercise 6.2. Let $q = 3$. Show that $L((0, 1), (1, 1))$ is the same line as $L((0, 1), (2, 2))$ and $L((1, 2), (1, 1))$. What is going on here?

Exercise 6.3. We have seen in the previous exercise that given $x \in \mathbb{F}_q^d$, there may exist $v \neq v' \neq (0, \ldots, 0)$ such that $L(x, v)$ and $L(x, v')$ are the same line. You probably observed that this problem takes

place if and only if $v = \lambda v'$ for some $\lambda \in \mathbb{F}_q$. (If you did not observe this, please verify it right away). Let V be a subset of $\mathbb{F}_q^d$ with the following properties:

- If $v \in V$, $v \neq (0, \ldots, 0)$.
- If $v \in V$ and $v' \in V$, there does not exist $\lambda \in \mathbb{F}_q$ such that $v = \lambda v'$.

Suppose that V is *maximal* in the sense that it is impossible to even a single element to V without violating one of the properties stated above. (Note that without this restriction we may simply take V to consist of a single point, say, $(1, 0, \ldots, 0)$).

Compute $\#V$, the number of elements of V.

In the standard (Euclidean) space, two different lines either do not intersect at all or intersect at a single point. We shall now verify that the same is true of lines in $\mathbb{F}_q^d$.

Exercise 6.4. Two different lines $L(x, v)$ and $L(x', v')$ in $\mathbb{F}_q^d$ either do not intersect at all or intersect at a single point. If $d = 2$, prove that two distinct lines $L(x, v)$ and $L(x', v')$ do not intersect if and only if there exists $\lambda \in \mathbb{F}_q$ such that $v = \lambda v'$.

We are now ready to state the main problem to be explored in this chapter.

Conjecture 6.4. *Let $B \subset \mathbb{F}_q^d$, $d \geq 2$, such that for every $v \in \mathbb{F}_q^d$ with $v \neq (0, \ldots, 0)$, there exists $x \in \mathbb{F}_q^d$ so that $L(x, v) \subset B$. Then there exists $C > 0$, independent of q, such that*

$$\#B \geq Cq^d.$$

To put it simply, the Besicovitch/Kakeya conjecture says that if B contains a line with every possible slope, then this set occupies a positive proportion of the points in $\mathbb{F}_q^d$. In short,

MANY SLOPES $\rightarrow$ MANY POINTS

The Besicovitch/Kakeya conjecture is solved in two dimensions. However, in higher dimensions, it is far from being resolved. For example, the best known result in three dimensions (see [**14**]) is that

$$\#B \geq Cq^{\frac{5}{2}+10^{-10}}.$$

One of the motivations behind this book is to convince you to dive head first into this mysterious problem which is not terribly likely to be completely solved any time soon.

1. Notes, remarks and difficult questions

There is an issue that we quite intentionally brushed under the rug until now. Did we ever use the fact that $\mathbb{F}_q$ is a field? Did we ever prove that $\mathbb{F}_q$ is a field? Is the conclusion of Exercise 6.4 still true if $\mathbb{F}_q$ is not a field? The answers are YES, NO and NO...

How do we prove that $\mathbb{F}_q$ is a field? First observe that if a, b are non-zero elements in $\mathbb{F}_q$, then $ab \neq 0$, since otherwise we would produce a factorization of a prime number q. We are now ready to produce inverses. Given $a \neq 0$ in $\mathbb{F}_q$, consider

$$a,\ a^2,\ \ldots, a^{q-1}. \tag{6.5}$$

If all these are distinct, there exists k such that $a^k = 1$ since none of the powers of a can be 0 by the observation above. This would mean that $a \cdot a^{k-1} = 1$, so a^{k-1} is the multiplicative inverse of a. The only obstacle to this parade is that the entries in (6.5) may not be distinct. In that case $a^k = a^j$ for some $k > j$. But then $a^{k-j} = 1$ and we are done anyways.

The rest of the field properties are easy to check. Do it!

Exercise 6.5. I claim above that the conclusion of the Exercise 6.4 no longer holds if $\mathbb{F}_q$ is not a field, which would be the case if q is not prime. Can you produce an explicit counter-example?

More importantly, does the Besicovitch-Kakeya problem still make sense if $\mathbb{F}_q$ is not a field. Elaborate...

Exercise 6.6. Let q be an odd prime and define

$$F = \mathbb{F}_q[\sqrt{2}] = \{a + b\sqrt{2} : a, b \in \mathbb{F}_q\}.$$

Prove that F is a field by explicitly exhibiting inverses and things of that nature. Go through all the exercises in the chapter with $\mathbb{F}_q$ replaced by F and convince yourself that everything still works and makes sense.

Chapter 7

Besicovitch-Kakeya conjecture in two dimensions

In the previous chapter we have innocently introduced a conjecture which basically says that if a subset of $\mathbb{F}_q^d$ contains a line in every "direction" then this subset contains a positive proportion of all the points in $\mathbb{F}_q^d$. What it says is that small sets do not contain roads going

Figure 1. Artistic depiction of a Besicovitch-Kakeya set

in all, or even most, directions. This is far from apparent because roads can intersect often, potentially making the set much smaller!

In this chapter we verify Conjecture 6.4 in the case $d = 2$. Once again, no special background is required. The reader is encouraged to check out the information related to the Besicovitch/Kakeya problem on the web page of Izabella Laba (www.math.ubc.ca), Terry Tao (www.math.ucla.edu), and the author (www.math.missouri.edu), as well as the original papers by Katz and Tao ([**12**]), Katz, Laba and Tao ([**14**]) and Wolff ([**21**]).

What you should be asking yourselves at every step is where are we using the peculiarities of the two-dimensional space, and why this approach should be harder in higher dimensions.

We have a set $B \subset \mathbb{F}_q^2$ which contains a line in every direction. This means that there exist lines $L_1, L_2, \ldots, L_{q+1}$ entirely contained in B with the additional property that any pair of these lines intersects at exactly one point. How do we know this? An obnoxious answer is that you verified exactly this in the exercises at the end of the previous chapter. Let's discuss it again, however. Consider $L(x, v)$ in two dimensions. How many choices are there for v? Well, $v = (v_1, v_2)$, so there are $q^2 - 1$ choices, since $v = (0, 0)$ is forbidden. On the other hand, multiplying v by $\lambda \in \mathbb{F}_q$ leads to the same line. How many λs are there? Since it makes no sense to use $\lambda = 0$, there are $q - 1$ relevant λs. It follows that B indeed contains $\frac{q^2-1}{q-1} = q + 1$ lines with different "slopes". By Exercise 6.4 each pair of such lines intersects at exactly one point.

Before we get on with the precise calculations, let's try to understand why the Besicovitch-Kakeya conjecture should be true in two-dimensions. As we have just seen, B contains $q + 1$ lines of different "slopes". Choose one of these lines and call it the stem. The other q lines intersect this stem forming a sort of hairbrush. Since two of these lines intersect at exactly one point, it is pretty clear that the total number of points in B is at least $(q + 1) \cdot \frac{q}{2} = \frac{q(q+1)}{2}$. Of course, we need to make this argument precise, which is what we are about to do.

All the tools are now in place. Let $B' = \cup_{i=1}^{q+1} L_i$. Since $B' \subset B$, it suffices to prove that $\#B' \geq Cq^2$. Let $\chi_{L_i}(x) = 1$ if $x \in L_i$ and

0 otherwise. We must somehow take advantage of the fact that we have $q+1$ lines with each pair intersecting at exactly one point. How do we "encode" intersections? Well,

$$\sum_{x\in B'} \chi_{L_i}(x)\chi_{L_j}(x) = \#\{x \in B' : x \in L_i \text{ and } x \in L_j\} = \#(L_i \cap L_j),$$

since $L_i \subset B'$ and $L_j \subset B'$ by assumption. With this observation in tow, consider

$$\sum_{x\in B'} [\chi_{L_1}(x) + \cdots + \chi_{L_{q+1}}(x)]^2 = \sum_{x\in B'} \sum_{i=1}^{q+1}\sum_{j=1}^{q+1} \chi_{L_i}(x)\chi_{L_j}(x) \tag{7.1}$$

$$= \sum_{i=1}^{q+1}\sum_{j=1}^{q+1} \#(L_i \cap L_j) = 2q(q+1),$$

where to obtain the second line we used the fact that $\#(L_i \cap L_j) = 1$ if $i \neq j$.

All of this is very nice, but we need to somehow get a hold on $\#B'$. This is where the first chapter of this book comes in handy. The left hand side of the first line of (7.1) is a sum of something squared. This immediately :) reminds us of the Cauchy-Schwarz inequality! Indeed, C-S tells us that

$$\left(\sum_{x\in B'} [\chi_{L_1}(x) + \cdots + \chi_{L_{q+1}}(x)] \cdot 1\right)^2$$

$$\leq \#B' \cdot \sum_{x\in B'} [\chi_{L_1}(x) + \cdots + \chi_{L_{q+1}}(x)]^2.$$

Plugging in the right hand side of (7.1) we see that

$$\left(\sum_{x\in B'} [\chi_{L_1}(x) + \cdots + \chi_{L_{q+1}}(x)]\right)^2 \leq \#B' \cdot 2q(q+1),$$

or, equivalently,

$$\#B' \geq \frac{\left(\sum_{x\in B'} [\chi_{L_1}(x) + \cdots + \chi_{L_{q+1}}(x)]\right)^2}{2q(q+1)}. \tag{7.2}$$

We seem to be getting somewhere provided we can evaluate the numerator of (7.2). We have

$$\sum_{x\in B'}[\chi_{L_1}(x)+\cdots+\chi_{L_{q+1}}(x)]=\sum_{x\in B'}\sum_{j=1}^{q+1}\chi_{L_j}(x)=\sum_{j=1}^{q+1}\sum_{x\in B'}\chi_{L_j}(x).$$

Since $L_j\subset B'$,

$$\sum_{x\in B'}\chi_{L_j}(x)=\#L_j=q.$$

We conclude that

$$\left(\sum_{x\in B'}[\chi_{L_1}(x)+\cdots+\chi_{L_{q+1}}(x)]\right)^2=q^2(q+1)^2.$$

Plugging this into (7.2) yields

$$\#B'\geq\frac{q(q+1)}{2},$$

as promised.

The result we just presented was first proved in $\mathbb{R}^2$ by R. Davies in 1971 ([**5**]), though the proof above is much closer to the one given for a related problem by A. Cordoba ([**4**]).

This seems to be the end of the story in two dimensions. Unfortunately, (or rather fortunately) mathematicians always find a way to complicate things. The following exercises give a taste of things to come in Part II of these notes where the level of fun (and pain) will get wrenched up another notch.

Exercise 7.1. Find the smallest possible Besicovitch/Kakeya subset of $\mathbb{F}_q^2$. We know that it contains at least $\frac{q(q+1)}{2}$ elements. Get as close to this number as you can. Hint: consider $S=\{(x_1,x_2)\in\mathbb{F}_q^2:$ $x_1+x_2^2$ is a square in $\mathbb{F}_q\}$. (A number s is a square in $\mathbb{F}_q$ if there exists $u\in\mathbb{F}_q$ such that $s=u^2$ in the world of $\mathbb{F}_q$).

1. Notes, remarks and difficult questions

Let $0<\alpha<1$. Suppose that we only assume that B is a subset of $\mathbb{F}_q^2$ with the property that for every $v\neq(0,0)$, $v\in\mathbb{F}_q^2$, there exists an $x\in\mathbb{F}_q^2$ such that more than q^α points of $L(x,v)$ are contained in B'.

Exercise 7.2. What can you say about $\#B$? Once you obtain an answer, try to determine whether your estimate is "reasonable". More precisely, for various values of $\alpha < 1$, experiment with constructions of subsets of $\mathbb{F}_q^2$ satisfying the required properties. This type of a formulation of the Kakeya problem is due to Hillel Furstenberg. See [**12**] and the references contained therein.

Chapter 8

A gentle entry into higher dimensions

Higher dimensional space is very annoying. It is no longer true that two lines are either "parallel" or intersect at a single point. It is quite easy for two lines to simply be in "parallel" planes which makes the structure of Besicovitch/Kakeya sets much harder to understand.

8.1. Bourgain's bush argument (late 80s). In this section we abandon our policy of systematically referencing the results we present. Instead, we refer the reader to Tom Wolff's beautiful survey article ([**21**]) where all the relevant references are present.

Figure 1. A less dramatic depiction of a Besicovitch-Kakeya set

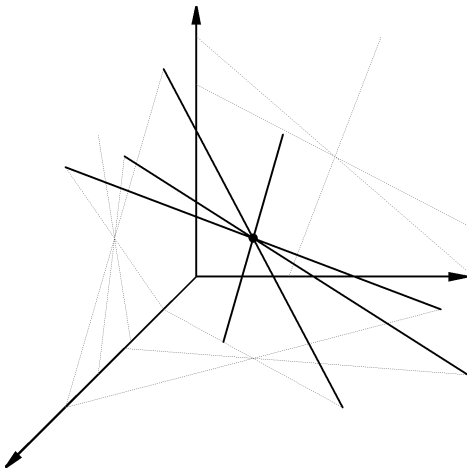


Figure 2. Bourgain's bush

Let's start with the following simple observation. Let B be a Besicovitch/Kakeya set in $\mathbb{F}_q^d$. How many lines must this set contain? Well, we know (from where?) that the answer is $\approx q^{d-1}$ (I am being obnoxious again...). Let's see why that is. Consider a line $L(x,v)$. We have q^d-1 choices for v since $v=(0,\ldots,0)$ is not allowed. As before, v and $v'=\lambda v$, $\lambda\in\mathbb{F}_q$, $\lambda\neq 0$, lead to the same line. It follows that the number of distinct lines in B is at least $\frac{q^d-1}{q-1}\geq\frac{q^{d-1}}{2}$.

Suppose that $\#B\leq\frac{q^{\frac{d+1}{2}}}{4}$. You might and should be wondering where this number is coming from. The honest answer is that this number is here because it makes the argument below work out. However, we could have just as easily kept this number a mystery and then plugged it in at the end when its value becomes clear. When you are done reading this proof, I urge you to go through the argument again and run it with a "variable" upper bound for $\#B$ and convince yourselves that the choice above is forced upon you... Moving right along, if we do assume that $\#B\leq\frac{q^{\frac{d+1}{2}}}{4}$, then at least one point of B must lie on at least

$$L=\frac{q^{\frac{d-1}{2}}}{2} \tag{8.1}$$

lines entirely contained in B. To see this, first observe the number of pairs (p, l), where $p \in B$, l is a line contained in B, and p lies on l, is at least $q \cdot \frac{q^{d-1}}{2} = \frac{q^d}{2}$ by the argument in the previous paragraph. By assumption, the number of points in B is $\leq \frac{q^{\frac{d+1}{2}}}{4}$. Then (8.1) follows since

$$\frac{q^d}{2} \leq \# \cup_p \cup_l \{(p, l) : p \in l\} \leq \frac{q^{\frac{d+1}{2}}}{4} \max_p \#\{(p, l) : p \in l\},$$

where p is a point in B and l is a line in B. It follows that

$$\max_p \#\{(p, l) : p \in l\} \geq \frac{q^{\frac{d-1}{2}}}{4},$$

as advertised. What we just proved is that there exists a point $p_0 \in B$ which belongs to at least $\frac{q^{\frac{d-1}{2}}}{4}$ lines in B. Since each of these lines contains $q - 1$ points aside from p_0,

$$\#B \geq 1 + L(q - 1) \geq \frac{q^{\frac{d+1}{2}}}{4}.$$

Thus we have shown that Besicovitch/Kakeya sets in $\mathbb{F}_q^d$ are $\frac{d+1}{2}$ "dimensional". This is horribly unsatisfactory since our goal is d, not $\frac{d+1}{2}$, and we can already do better than $\frac{d+1}{2}$ when $d = 2$. We did take an important step in the right direction, though, as the techniques we just developed will come in handy in a moment.

8.2. Wolff's hairbrush argument (mid 90s). What was the essence of the bush argument? If lines do not intersect much, we win because there are points all over the place. If lines do intersect, we look for places where lots of lines intersect in the same place. We call such a happy meeting place a bush. What we did above is first argued that there must exist a fairly large bush. We then estimated the number of points in this bush and obtained the desired estimate. As cute as this argument is, it is hopelessly naive if we are to get anywhere close to the full Besicovitch/Kakeya conjecture.

Our next step in the direction of fame and glory (don't get too excited) is the hairbrush construction. Let B be a Besicovitch/Kakeya set and suppose that $\#B \leq q^{\frac{d+2}{2}}$. We repeat the argument we used in the two-dimensional case. We know by above that B contains at

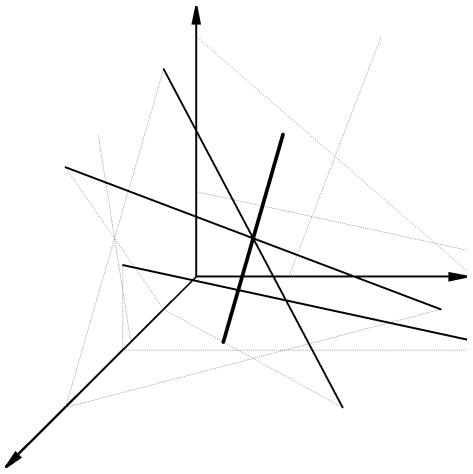


Figure 3. Wolff's hairbrush

least $\frac{q^{d-1}}{2}$ lines with distinct "slopes". Let B' be the union of these lines. Reusing the proof of the two-dimensional Besicovitch/Kakeya conjecture, we have

$$q^2k^2 = \left(\sum_{x\in B'} [\chi_{L_1}(x) + \cdots + \chi_{L_k}(x)]\right)^2 \tag{8.2}$$

$$\leq \#B' \cdot \sum_{x\in B'} [\chi_{L_1}(x) + \cdots + \chi_{L_{q+1}}(x)]^2$$

$$= \#B' \cdot \sum_{x\in B'} \sum_{i=1}^{k} \sum_{j=1}^{k} \chi_{L_i}(x)\chi_{L_j}(x)$$

$$= \#B' \cdot \sum_{i=1}^{k} \sum_{j=1}^{k} \#(L_i \cap L_j),$$

where k is the number of lines (which by above is $\geq \frac{q^{d-1}}{2}$).

Exercise 8.1. Why is the first line in (8.2) true? Did we use an inequality with a name in the second line? Which one?

Since we have assumed that $\#B \leq q^{\frac{d+2}{2}}$, we also have $\#B' \leq q^{\frac{d+2}{2}}$. Plugging this into (8.2) we get

$$\sum_{i=1}^{k} \sum_{j=1}^{k} \#(L_i \cap L_j) \geq \frac{q^{\frac{3d-2}{2}}}{2}.$$

It follows that there exists $i = i_0$ such that

$$\sum_{1 \leq j \leq k; j \neq i_0} \#(L_j \cap L_{i_0}) \geq \frac{\frac{q^{\frac{3d-2}{2}}}{2}}{\frac{q^{d-1}}{2}} - q \geq \frac{q^{\frac{d}{2}}}{2}.$$

We just proved that there exists a line L_{i_0}, called the base of the hairbrush, such that at least $m = \frac{q^{\frac{d}{2}}}{2}$ other lines contained in B' intersect it. We call this collection of m lines the hairbrush, denoted by H.

Let Π_j denote the two-plane determined by L_{i_0} and L_j. Suppose that Π_j contains $n_j \geq 1$ lines from the hairbrush. We want to estimate $\#(\Pi_j \cap H)$ from below. Since Π_j is a two-dimensional plane, we should be able to use the solution of the two-dimensional Besicovitch-Kakeya conjecture. Unfortunately, in that section we only learned to deal with sets containing approximately q lines with different slopes. In this case we have n_j lines, which may be smaller than q. This predicament forces us to rewrite the argument in Chapter 7 for the purpose at hand. Let $L_1, \ldots, L_{n_j}$ be the lines in the hairbrush (after possibly doing some relabelling) that are contained in $\Pi_j \cap H$. We have

$$q^2(n_j+1)^2 = \left(\sum_{x \in \Pi_j \cap H} [\chi_{L_1}(x) + \cdots + \chi_{L_{n_j}}(x)] \right)^2$$

$$\leq \#(\Pi_j \cap H) \sum_{x \in \Pi_j \cap H} [\chi_{L_1}(x) + \cdots + \chi_{L_{n_j}}(x)]^2$$

$$= \#(\Pi_j \cap H) \sum_{i=1}^{n_j+1} \sum_{i'=1}^{n_j+1} \#(L_i \cap L_{i'})$$

$$= \#(\Pi_j \cap H) \cdot ((n_j+1)q + (n_j+1)(n_j)) \leq 2 \cdot \#(\Pi_j \cap H) \cdot (n_j+1)q.$$

It follows that

$$\#(\Pi_j \cap H) \geq \frac{1}{2}(n_j + 1)q \geq \frac{qn_j}{4}.$$

Exercise 8.2. Why does $n_j + 1$ appear all over the place instead of n_j? Hint: Don't forget the base of the hairbrush...

We are almost done since

$$\#B \geq \#H \geq \frac{q}{4}\sum_{j=1}^{t} n_j = \frac{qm}{4} \geq \frac{q^{\frac{d+2}{2}}}{8}.$$

We just proved that if B is a Besicovitch/Kakeya set in $\mathbb{F}_q^d$, then

$$\#B \geq \frac{q^{\frac{d+2}{2}}}{8}.$$

This is not quite the Besicovitch/Kakeya conjecture, but we are getting closer!

1. Notes, remarks and difficult questions

Does a hairbrush in the argument above need to contain a line of every "slope"?

Exercise 8.3. Given an explicit example proving that it does not. Now suppose that you have a Besicovitch/Kakeya set containing a hairbrush containing a line with every possible "slope". Prove that $\#B \geq Cq^d$. Prove that the same conclusion follows if instead of assuming that the hairbrush contains a line with every possible "slope", it only contains at least cq^{d-1} lines with different "slopes".

Can you construct an example of a Besicovitch/Kakeya set B such that no hairbrush contains at least cq^{d-1} lines with different "slopes"?

An approach used in the aforementioned paper of Katz, Laba and Tao to improve the $\frac{5}{2}$ exponent in three dimensions is to prove that if the size of the Besicovitch set really is $\approx q^{\frac{5}{2}}$, then the Besicovitch set in question must basically consist of copies of hairbrushes in a fairly regular arrangement. Would you like to take a shot at this type of an argument in vector spaces over finite fields?

Chapter 9

Some basic counting, probability and a few twists

In this chapter we explore some basic counting techniques and elementary probability. The treatment is far from systematic because we hope to provide new perspectives on probability and to further close the mythical divide between "discrete" and "continuous" mathematics. For those with reasonable familiarity with discrete counting and probability, this chapter will serve as a springboard of sorts to the more complicated probabilistic arguments in the upcoming chapters. For a thorough treatment of basic probability, the reader is encouraged to look at a beautiful book by Feller ([**7**]), entitled "Probability theory and its applications".

9.1. Factorials and choosing. Steve, Gina and Alex run a hundred yard dash. How many different outcomes are possible? Observe that we are not asking how many different outcomes are likely... The choices are SGA, SAG, GSA, GAS, ASG, and AGS- six in total. Can we do something a little more efficient than listing all the possibilities? Of course we can. We have three choices for the first slot in the race. Once the first slot is filled, there are only two choices left for the second slot, and once the second slot is filled, the third slot is

determined. This means that the total number of possibilities is $1 \cdot 2 \cdot 3 = 6$. In general, we have the following result.

Theorem 9.1. *The number of ways of ordering n objects is $n! = 1 \cdot 2 \ldots\ldots n$.*

We can prove this result by the same argument with which we handled the case $n = 3$, but we will use this opportunity to review mathematical induction. The case $n = 1$ is clear. Let us assume that the theorem holds for $n - 1$. Take one of the objects from the group of n. Its number in the ordering is between 1 and n, inclusive. This means that this object has one of n possibilities in the order. The remaining $n - 1$ objects can be ordered in $(n - 1)!$ ways by the induction hypothesis. This means that the total number of ways of ordering n objects is $(n - i)! \cdot n = n!$ and we are done.

Let us complicate things a bit. Suppose that 3 soldiers need to be chosen for a dangerous mission out of a platoon of twenty stationed in Iraq. How many ways are there to do that? Well, there are twenty ways to choose the first soldier. Once the first soldier is chosen, there are nineteen ways of choosing the second soldier, and, similarly, eighteen ways of choosing the third soldier. It appears therefore that there are $20 \cdot 19 \cdot 18$ ways of choosing the squad of three. Unfortunately, we are making a rather common and silly mistake. The $20 \cdot 19 \cdot 18$ counting actually distinguishes between, say, the Joe, John, Greg, and Greg, Joe, John combinations. I think you would agree that Joe, John and Greg do not particularly care in what order they were chosen for the mission, only the fact that they were in fact chosen for it! This means that the actual count is $20 \cdot 19 \cdot 18$ divided by the number of ways of ordering three soldiers. By Theorem 9.1, there are $3! = 6$ ways of ordering three soldiers, so the number of ways of choosing the group of three for the dangerous mission is $\frac{20 \cdot 19 \cdot 18}{6}$ The general result is the following.

Theorem 9.2. *The number of ways of choosing k indistinguishable objects out of n $(1 \leq k \leq n)$ is $C(k, n) = \frac{n!}{k!(n-k)!}$.*

The proof is just a generalization of the argument in the previous paragraph. We have n choices for the first of k objects, $n - 1$ for the second, $n - 2$ for the third, and, finally, $n - k + 1$ for the kth. On

the other hand, these k victims do not care about the order in which they are chosen any more than the guys in the example above, so the number of possibilities is indeed $\frac{(n(n-1)\ldots(n-k+1)}{k!} = C(k,n)$. We are smoking!

9.2. Binomial theorem and subsets of finite sets. What else can we do? The possibilities are certainly endless. Instead of counting, we are going to apply the techniques we have explored so far to do something useful. What is $(a+b)^n$ where n is a positive integer? If $n = 1$, the answer is clear enough. If $n = 2$, we have $(a+b)^2 = (a+b)(a-b) = a^2 + 2ab + b^2$. Fantastic... We may even recall that $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$. The idea is to not have to recall anything, so what is really going here? We have

$$(a+b)^n = (a+b)\cdot(a+b)\ldots(a+b).$$

How do actually multiply this quantity out? From each set of parentheses we choose either an a or a b. Note that we have exactly two choices in each case. We thus end up with $a^k b^{n-k}$ since the powers must add up to n. How many times does the term $a^k b^{n-k}$ arise? Observe that once we know which sets of parentheses a's come from, we know exactly where the b's come from. This means that the term $a^k b^{n-k}$ arises exactly as many times as there are ways of choosing k objects out of n. We have just deduced the celebrated binomial theorem.

Theorem 9.3. *Let a and b be real numbers. Then*

$$(a+b)^n = \sum_{k=0}^{n} C(k,n) a^k b^{n-k}.$$

DO NOT stop here! Always look for generalizations and variants! ALWAYS!! Yes, I am shouting... What about $(a+b+c)^n$, $(a+b+c+d)^n$, ..., $(a_1+a_2+\cdots+a_N)^n$? WORK IT OUT!!! If you get stuck, though, move on and come back to it later. The train of progress must keep rolling...

Theorem 9.3 has a beautiful corollary which is worth exploring in its own right. Let $a = 1$ and $b = 1$. Then it follows from Theorem 9.3 that

$$2^n = \sum_{k=0}^{n} C(k,n). \tag{9.4}$$

Let us verify this fact directly by using another kind of counting. Let S be a set with n elements. How many subsets does S have? Let us first consider an example. Suppose that $S = \{a, b\}$, a set with two elements. Its subsets are $\emptyset$, the empty set, $\{a\}$, $\{b\}$, and $\{a, b\}$. In total we have four different subsets. Let us now take $S = \{a, b, c\}$, a set with three elements. The subsets are $\emptyset$, $\{a\}$, $\{b\}$, $\{c\}$, $\{a, b\}$, $\{a, c\}$, $\{b, c\}$, and $\{a, b, c\}$. This time we have eight subsets in total. A pattern is developing, but, to paraphrase Sherlock Holmes, it is dangerous to draw conclusions in the absence of sufficient facts, so a more systematic approach is needed.

We now have a mechanism for counting the number of subsets. Given a subset of S, we walk up to each element of S and ask, very sternly, "are you or are you not a member of this subset?!". The answer is either YES or NO, 1 or 0. Thus each element has only two choices, and since there are n elements in total, the number of subsets of S is $2 \cdot 2 \ldots 2 = 2^n$! To put it another way, think of each subset of S as consisting of all n elements of S, but each element comes with an additional label indicating whether it is dead or alive. For example, if $S = \{a, b\}$, as above, the subset $\{a\}$ can be thought of as $\{a(\text{alive}), b(\text{dead})\}$. The counting is now easy, even in general. Every element is either alive or dead, and there are n elements, so the total number of possibilities is 2^n.

Let us now count the number of subsets of the set S with n elements in another way. How many subsets of S have k elements? Well, by Theorem 9.2 this number is exactly $C(k, n)$. It follows that the total number of subsets is $\sum_{k=0}^{n} C(k, n)$ and so the equation 9.4 is vindicated.

9.3. Informal introduction to expected value. As you have surely observed, and may even remember from various sources, we are using counting to verify identities and vice-versa! Let's do another one... Flip a fair coin. The probability of getting heads is $\frac{1}{2}$ and probability of getting tails is $\frac{1}{2}$, so the probability of getting k

heads (or k tails) after k flips is 2^{-k}. The question we ask is, how many flips should we expect to make before we get heads? Let us rephrase the problem slightly. Since we are asking for how many flips we "expect" to make, we are implicitly averaging over all the possible outcomes of our coin flips in order to determine one that is "expected". Another way of thinking about it is the following. Suppose that we live in the world where all possible outcomes take place and we must pay a portion of a dollar for one flip, a portion of two dollars for two flips, a portion of three dollars for three flips, and so on. Moreover, we only pay for bad outcomes, the ones where heads is not achieved. If you think about it for a moment, you will see that the amount of money we pay out in the end is precisely the expected number of flips because it is only the bad outcomes that cause the process to continue and the number of needed flips to increase! We shall henceforth refer to this idea as the pay-off analogy.

Now if we make k flips, the probability of the outcome where no heads are achieved is 2^{-k}. This means that the expected number of flips to get heads is $\sum_{k=0}^{\infty} \frac{k}{2^k}$. What is this number? Well, let us calculate the "expected" number of flips needed to get heads in another way. The probability of getting heads on the first try is $\frac{1}{2}$, so the "expected" number of heads after the first try is $\frac{1}{2}$, as paradoxical as this may sound. This means that the expected number of heads after the second flip is $\frac{1}{2}+\frac{1}{2}=1$, and we conclude that the "expected" number of flips needed to get heads is 2! This is exciting in itself, but we also deduce a beautiful identity

$$\sum_{k=0}^{\infty} \frac{k}{2^k} = 2. \tag{9.5}$$

While we are at it, let us compute $\sum_{k=1}^{\infty} \frac{k}{2^k}$ in another way without appealing to any (more or less) prior knowledge. Observe that $k = \sum_{j=1}^{k} 1$, so

$$\sum_{k=1}^{\infty} \frac{k}{2^k} = \sum_{k=0}^{\infty} \sum_{j=1}^{k} 2^{-k} \tag{9.6}$$

$$= \sum_{j=1}^{\infty} \sum_{k=j}^{\infty} 2^{-k}. \tag{9.7}$$

How did the last line follow from the previous one? Well, looking at the second line above, we have $1 \leq j \leq k$ and $1 \leq k < \infty$. This means that $1 \leq j \leq k < \infty$, which can be rewritten as $j \leq k < \infty$ and $1 \leq j < \infty$.

We must still compute the expression in (9.7). Define

$$s_j = \sum_{k=j}^{\infty} 2^{-k}. \tag{9.8}$$

We have

$$s_j = 2^{-j} + 2^{-j-1} + \ldots,$$

so

$$2^{-1} s_j = 2^{-j-1} + 2^{-j-2} + \ldots.$$

Subtracting the two expressions we see that

$$\frac{s_j}{2} = 2^{-j},$$

so

$$s_j = 2^{-j+1}. \tag{9.9}$$

Plugging (9.9) into (9.7) we see that

$$\sum_{j=1}^{\infty} \sum_{k=j}^{\infty} 2^{-k} = \sum_{j=1}^{\infty} 2^{-j+1} = 2 \cdot \sum_{j=1}^{\infty} 2^{-j}.$$

By the definition given in (9.8), this expression equals $2 \cdot s_1$, and by (9.9), $s_1 = 1$. It follows that

$$\sum_{k=1}^{\infty} \frac{k}{2^k} = 2,$$

as before. In other words, contradictions in foundations of mathematics have not yet been discovered in this book!

We are not finished being annoying ... Let us compute $\sum_{k=1}^{\infty} \frac{k}{2^k}$ in yet another way! Let

$$f(x) = \sum_{n=0}^{\infty} x^n.$$

We may or may not remember that if $|x| < 1$, this sum converges, and, in that range, $f(x)$ is a differentiable function of x. Differentiating term by term, we see that

$$f'(x) = \sum_{n=0}^{\infty} nx^{n-1} = \sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{x}\sum_{n=1}^{\infty} nx^n. \tag{9.10}$$

Let us get a formula for $f(x)$. We have just done it above in the case $x = \frac{1}{2}$, but we need the general case now. Not that it is any different, but repetition is the mother of progress... We have

$$f(x) = 1 + x + x^2 + \cdots + x^n + \ldots,$$

so

$$xf(x) = x + x^2 + \cdots + x^n + x^{n+1} + \ldots.$$

Subtracting we see that

$$f(x)(1-x) = 1,$$

so

$$f(x) = \frac{1}{1-x}.$$

Differentiating this formula we see that

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

Plugging this into (9.10) we see that

$$\frac{1}{(1-x)^2} = \frac{1}{x}\sum_{n=1}^{\infty} nx^n. \tag{9.11}$$

We are ready to strike! Let $x = \frac{1}{2}$ and plug it into (9.11). We obtain

$$4 = 2\sum_{n=1}^{\infty} \frac{n}{2^n},$$

and so

$$\sum_{n=1}^{\infty} \frac{n}{2^n} = 2$$

as before! Again, contradictions have been averted!

We have discovered a very nice theme and I want it to get firmly cemented in your brain, so let us do another cute example. Suppose

that a drunk hunter shoots at n raccoons. Since he is not exactly at his best, each raccoon has the probability p, $0 < p < 1$, of surviving the experience. What is the expected number of surviving raccoons at the end of the carnage? As before, we will proceed in two different ways. First, by simple counting. Suppose that there is only one raccoon present, what is the expected number of surviving raccoons after the shot? Well, the raccoon survives with probability p, so following our monetary pay-off analogy, the expected number of surviving raccoons will be p. As paradoxical as this sounds, consider the special case when $p = \frac{1}{2}$. Then the raccoon lives or dies with equal probability. If the lucky creature lives, we have one live raccoon. If the poor creature dies, we have one dead raccoon. The average is $\frac{1}{2}\ldots$ as we claimed... Nothing really changes if we have n raccoons instead of 1. The probabilities are independent, so the expected number of surviving raccoons is

$$\frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{2} = \frac{n}{2},$$

which means that half of the raccoons are expected to survive- not terribly shocking...

Let us now see what the cat drags in if we employ the pay-off analogy to the set of n raccoons instead of just one. Following the logic we developed for the case of coin flips, we must pay a portion of one dollar if one raccoon survives, a portion of two dollars if two raccoons survive, and so on, where the "portion" is the probability that EXACTLY k raccoons survive. What is this probability? Well, for k raccoons to survive, $n - k$ raccoons must die. The probability that a GIVEN SET OF k raccoons survive while $n - k$ raccoons die is

$$p^k \cdot (1-p)^{n-k}.$$

Since we are interested in ANY set of k raccoons, we must take into account the fact that k raccoons can be chosen in $C(k, n)$ different ways out of the set of n raccoons present at the carnage. It follows that the probability of EXACTLY k raccoons surviving is

$$C(k,n)p^k \cdot (1-p)^{n-k},$$

and the expected value of the number of surviving raccoons is

$$\sum_{k=0}^{n} k \cdot C(k,n)p^k(1-p)^{n-k}, \tag{9.12}$$

from which we deduce another attractive identity

$$E(p) = \sum_{k=0}^{n} k \cdot C(k,n)p^k(1-p)^{n-k} = np. \tag{9.13}$$

As you might have guessed, the suffering is not over! In order to be doubly sure, we are going to verify (9.13) by another calculation and, hopefully, learn a thing or two in the process.

Let us begin with a more innocent question. What is

$$f(p) = \sum_{k=0}^{n} C(k,n)p^k(1-p)^{n-k}? \tag{9.14}$$

Well, each term is the probability that EXACTLY k raccoons live. At the end of the carnage, some number of raccoons survive. We do not know how many, but some number of raccoons do survive! Even if "some" happens to be 0. It follows that the sum of these probabilities must be 1, and so the expression in (9.14), $f(p) = 1$! (Take a look at Theorem 9.3 as well...) Do you see what I am saying? The sum of probabilities of all possible events is always equal to 1 because SOMETHING happens with probability one!

How do we recover $E(p)$ from $f(p)$. In other words, how do we make a factor of k appear? Well, the derivative of p^k with respect to p is kp^{k-1}, so we have an idea! More precisely, we have

$$\begin{aligned} 0 = f'(p) &= \sum_{k=0}^{n} C(k,n)kp^{k-1}(1-p)^{n-k} \\ &\quad - \sum_{k=0}^{n} C(k,n)p^k(n-k)(1-p)^{n-k-1} \\ &= \sum_{k=0}^{n} C(k,n)kp^{k-1}(1-p)^{n-k} \\ &\quad - n\sum_{k=0}^{n} C(k,n)p^k(1-p)^{n-k-1} \end{aligned}$$

$$+\sum_{k=0}^{n} kC(k,n)p^k(1-p)^{n-k-1} = I + II + III.$$

Multiplying and dividing by p, we see that

$$I = \frac{1}{p}\sum_{k=0}^{n} kC(k,n)p^k(1-p)^{n-k} = \frac{1}{p}E(p).$$

Multiplying and dividing by $(1-p)$ we see that

$$II = -\frac{n}{1-p}\sum_{k=0}^{n} C(k,n)p^k(1-p)^{n-k}$$

$$= -\frac{n}{1-p}f(p) = -\frac{n}{1-p}.$$

Multiplying and dividing by $1-p$ we see that

$$III = \frac{1}{1-p}E(p).$$

It follows that

$$0 = I + II + III = E(p)\left(\frac{1}{p} + \frac{1}{1-p}\right) - \frac{n}{1-p}.$$

It follows that

$$E(p)\cdot\frac{1}{p(1-p)} = \frac{n}{1-p},$$

so

$$E(p) = np,$$

as advertised. Once again contradictions in foundations of mathematics have been avoided. What a relief!

9.4. Guiness, O'Reilly's, and the random walk. We shall start with an amusing situation where Alex, after having consumed a dozen pints of Guiness, can only move up and down one street. We shall assume that he can take one step forward or one step backward with equal probability. Staying in one place is not allowed in this model and neither is falling, face first or otherwise. The question we ask is, what is Alex's expected distance from the starting point after n steps? Naive intuition may (or may not) tells us that this distance is 0 since Alex is equally likely to move forwards or backwards, but

this is not the case. Remember that we are measuring the expected DISTANCE from the starting point, not average position.

Using the pay-off analogy, we must pay a portion of one dollar for Alex being a distance one from the origin, a portion of two dollars for Alex being a distance two from the origin, and so forth. The "portion" is the probability that Alex is EXACTLY the distance k from the origin after n steps. To calculate this probability observe that to be a distance k from the origin, Alex needs to take a steps forward and $n-a$ steps backwards such that $|a-(n-a)| = |2a-n| = k$. This means that Alex must take $\frac{n+k}{2}$ steps forwards or $\frac{n-k}{2}$ steps backwards. Observe that this is only possible if n and k have the same parity, which means that they are either both odd or both even. If n and k do have the same parity, the distance k from the origin can be achieved in

$$C\left(\frac{n+k}{2}, n\right) + C\left(\frac{n-k}{2}, n\right)$$

ways and thus the probability that Alex is a distance k away from the origin is

$$\frac{C\left(\frac{n+k}{2}, n\right) + C\left(\frac{n-k}{2}, n\right)}{2^n} = \frac{2C\left(\frac{n+k}{2}, n\right)}{2^n}$$

since the total number FORWARD/BACK configurations of length n is exactly 2^n, and

$$C(l, n) = C(n-l, n).$$

Using the pay-off analogy, the expected distance from the origin is

$$2 \cdot \sum_{k=0}^{n} \frac{kC\left(\frac{n+k}{2}, n\right)}{2^n}. \tag{9.15}$$

This sum is quite difficult... Let us do something a little more clever and avoid this sum altogether. Let $X_1, X_2, ..., X_n$ each take on values of $+1$ or -1 with equal probability. Such objects are called random variables. Notice that we are working with them without really defining them rigorously. This is very much done on purpose. I want to understand these objects on the functional level before seeing

any formal definitions. Anyways, let us get back to business! The place where Alex ends up can be computed as

$$X_1 + X_2 + \cdots + X_n.$$

What is the expected value of this random quantity? Let us take advantage of the fact that the expected value of X_j^2 is clearly 1 since this is the only value X_j^2 ever takes on! With this in mind, let us compute the expected value of

$$(X_1 + X_2 + \cdots + X_n)^2$$
$$= X_1^2 + \cdots + X_n^2 + \sum_{i \neq j} X_i X_j = I + II.$$

Since the expected value of X_j^2 is 1, the expected value of I is n. Now, what is the expected value of II? I claim that it is 0. Why? Well, X_i takes on values ± 1 with equal probability, and so does X_j. Since $i \neq j$ (different steps) and Alex is moving without a plan (randomly), $X_i X_j$ also takes on values ± 1 with equal probability, so the expected value of $X_i X_j$ must also be 0. It follows that the expected value of II is 0 and thus the expected value of $(X_1 + \cdots + X_n)^2$ is n. We conclude that the expected value of $X_1 + \cdots + X_n$ is $\sqrt{n}$. It follows that the value of the sum in (9.15) is $\sqrt{n}$. This calculation saved us a lot of grief because, as we note above, calculating (9.15) directly is not in our plans for the moment. But, this book is far from being concluded...

9.4.1. *Is Alex likely to return?* An amusing but important question is whether random wanderings are likely to cause Alex to eventually end up in the same spot where she started. To put it another way, what is the probability that Alex will end up by the door of O'Reilly's Grill 15? Well, what is the probability that Alex's distance from the origin is 0 after n steps. Actually, let us take $2n$ steps to deal with the fact that after an odd number of steps one cannot possibly end up back at the origin. This probability is

$$p_0(n) = 2 \cdot \frac{C(n, 2n)}{2^{2n}},$$

so the question reduces to understanding how $C(n, 2n)$ behaves as n gets larger and larger. Or is it? Later in the notes we will analyze

$C(k,n)$ using more advanced methods, but recall that the purpose of these notes is not zoological exploration but rather to build up intuition and the courage to experiment. So let us just go for it!

What is the plan? Suppose that we can show that

$$\sum_{n=1}^{\infty} p_0(n) \tag{9.16}$$

diverges. We can then invoke the pay-off analogy yet again and argue that $p_0(n)$ is precisely the expected number of returns to the origin. The fact that the sum in (9.16) diverges would mean that Alex is likely to return to O'Reilly infinitely often!

We must still prove that the sum in (9.16) in fact diverges. We shall do it in a slightly inefficient way, but one that is "good enough".

Lemma 9.17. *Let n be a positive integer. Then*

$$\frac{C(n,2n)}{2^{2n}} \geq \frac{1}{2n}. \tag{9.18}$$

The statement is true for $n = 1$ by inspection. Suppose that the statement is true for n and we verify it for $n+1$. By definition of $C(n,2n)$ and the induction hypothesis,

$$\begin{aligned}\frac{C(n+1,2n+2)}{2^{2n+2}} &= \frac{C(n,2n)}{2^{2n}}\frac{(2n+1)(2n+2)}{4n^2} \\ &\geq \frac{1}{2n}\frac{(2n+1)(2n+2)}{4n^2},\end{aligned}$$

so we must show that

$$\frac{1}{2n}\frac{(2n+1)(2n+2)}{4n^2} \geq \frac{1}{2(n+1)},$$

which amounts to showing that

$$(2n+1)(n+1)^2 \geq 2n^3,$$

which is true by expansion.

The assertion that the sum in (9.16) diverges now follows from (9.18), the comparison test for sums, and the fact that $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges.

We conclude that Alex is indeed likely to come back to O'Reilly's infinitely often, even if he stumbles around randomly. To be fair, we

did assume that Alex moves in a straight line which is far from the truth. However, later in this book we shall see that even if Alex moves in a two-dimensional plane, also not exactly true, he is still likely to come back to O'Reilly's infinitely often...

There is some bad news, however. We proved that $\frac{C(n,2n)}{2^{2n}} \geq \frac{1}{2n}$, which is true and good enough for what we are doing, but it is not optimal. If we apply this bound to the random walk in two dimensions, we see that

$$\left(\frac{C(n,2n)}{2^{2n}}\right)^2 \geq \frac{1}{4n^2}.$$

The infinite sum of the quantity on the right converges, whereas the random walk in two dimensions still returns to the origin infinitely often! The truth of the matter is that

$$\frac{C(n,2n)}{2^{2n}} \approx \frac{1}{\sqrt{n}},$$

but this is much more difficult to prove. Give it a try!

9.5. Some continuous probability. Alex is drunk again, but this time he is stumbling around randomly inside a disk of radius one centered at the origin with equal probability of being at any given point inside. What is the probability that Alex is above the x-axis? Intuitively, the answer is $\frac{1}{2}$. However, unlike the discrete case where intuition fairly easily translates into rigor, at least on some level, things are a bit trickier now and require some care. Let S be a Riemann measurable subset of the disk. If you do not remember what that is, do not worry about it. Just think of D as a set over which you can integrate—that's all! We define the probability of being in D by the equation

$$p(S) = \int_D \chi_S(x,y)dxdy, \tag{9.19}$$

where D is the unit disk and $\chi_S(x,y) = 1$ if $(x,y) \in S$ and 0 otherwise. In other words

$$p(S) = \frac{\text{Area } (S)}{\text{Area } (D)} = \frac{\text{Area } (S)}{\pi}.$$

Therefore, if S is the portion of the disk that lives above the x-axis, its area is $\frac{\pi}{2}$, so $p(S) = \frac{\frac{\pi}{2}}{\pi} = \frac{1}{2}$.

9.5.1. *Sometimes it is not clear what probability means.* Let us stick with the disk of radius 1 centered at the origin. The boundary of this disk is the unit circle. Suppose that we choose an arc of this circle at random. What is the expected value of its length? Keep in mind that we never actually defined the expected value rigorously, but we do have that pay-off analogy. What does it mean in this context? The whole point is what exactly does *probability* mean? How do we choose an arc at "random"?

Possibility One: Parameterize the circle by its arc-length and denote the endpoints of the arc by angles θ and ϕ. The length of the arc is $|\theta - \phi|$. The expected value of the arc-length is then

$$E = \frac{1}{2\pi} \cdot \frac{1}{2\pi} \int_0^{2\pi} \int_0^{2\pi} |\theta - \phi| d\theta d\phi.$$

Why are we dividing by 2π? Because we are implicitly putting a probability function on the unit circle. Given a subset S of the unit circle, we define its probability, $p(S)$, in analogy with the case of the disk above, but the formula

$$p(S) = \frac{1}{2\pi} \int_0^{2\pi} \chi_S(\cos(\theta), \sin(\theta)) d\theta,$$

where $\chi_S(x, y)$ equals 1 if $(x, y) \in S$ and 0 otherwise. We still have not answered the question of why we are dividing by 2π. Well, suppose that S is the whole circle. We want its probability to be 1, right?! I mean, what is the probability that a point in the circle is in the circle? It is one, of course... This will only happen if we divide by 2π because the length of the circle of radius 1 is 2π.

Observe that the notion of expected value is more transparent all of a sudden. The expected value of the length is just the honest to goodness average of the length. Straightforward, right? We will see about that... Let us complete the calculation first. We have

$$E = 2 \cdot \frac{1}{2\pi} \cdot \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\theta} (\theta - \phi) d\phi d\theta$$

$$= 2 \cdot \frac{1}{2\pi} \cdot \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\theta} \theta d\phi d\theta$$
$$-2 \cdot \frac{1}{2\pi} \cdot \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\theta} \phi d\phi d\theta = I + II.$$

Now,

$$I = 2 \cdot \frac{1}{2\pi} \cdot \frac{1}{2\pi} \int_0^{2\pi} \theta^2 d\theta = 2 \cdot \frac{1}{2\pi} \cdot \frac{1}{2\pi} \frac{(2\pi)^3}{3} = \frac{4\pi}{3},$$

and

$$II = -2 \cdot \frac{1}{2\pi} \cdot \frac{1}{2\pi} \int_0^{2\pi} \frac{\theta^2}{2} d\theta = -\frac{2\pi}{3}.$$

It follows that

$$E = \frac{4\pi}{3} - \frac{2\pi}{3} = \frac{2\pi}{3}.$$

Possibility Two: Instead of defining a probability function on the unit circle, denoted by S^1, we shall define the probability measure on the whole disk. Let $A \in S^1 \times S^1$, where

$$S^1 \times S^1 = \left\{(u, v) : u \in S^1, v \in S^1\right\}.$$

Why are we talking about $S^1 \times S^1$? We have already implicitly dealt with it in Possibility One. The point is that each endpoint of the arc is on the circle S^1, so a pair of endpoints is on the product set $S^1 \times S^1$. Define the probability of A by the formula

$$p(A) = \frac{1}{\pi} \int \int_D \chi_{m(A)}(x, y) dx dy,$$

where D, as before, is the unit disk, and $m(A)$ is the image of A under the function m defined as follows. Since A is the subset of $S^1 \times S^1$, its points are of the form (u, v) where $u \in S^1$ and $v \in S^1$. Draw a straight line from u to v. Now draw a straight line from the center of the unit disk until it hits this line in the perpendicular fashion. The point of intersection is $m(u, v)$. Again, we divide by π to ensure that if A is all of $S^1 \times S^1$ then $p(A) = 1$ as it should.

Let us now compute the expected length of a "random" circular arc with respect to this new probability function. A bit of trigonometry (do it!) shows that if $(x, y) \in D$, the length of the corresponding arc is

$$2 \arccos(\sqrt{x^2 + y^2}),$$

and the expected value of the length is

$$\frac{2}{\pi}\int\int_D \arccos(\sqrt{x^2+y^2})dxdy = 4\int_0^1 \arccos(r)rdr.$$

Let us now watch Alex integrate by parts. We get

$$4\int_0^1 \arccos(r)rdr = 4\arccos(r)\frac{r^2}{2}\,|_0^1 + 4\cdot\frac{1}{2}\int_0^1 \frac{r^2}{\sqrt{1-r^2}}dr$$

$$= 2\int_0^1 \frac{r^2}{\sqrt{1-r^2}}dr.$$

Let $r = \sin(\theta)$, $dr = \cos(\theta)d\theta$. We get

$$2\int_0^{\frac{\pi}{2}} \sin^2(\theta)d\theta. \tag{9.20}$$

How do we do this? Recall that

$$\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta) = 1 - 2\sin^2(\theta),$$

so

$$\sin^2(\theta) = \frac{1-\cos(2\theta)}{2}.$$

BY THE WAY: If you do not remember how to derive these trigonometric identities, figure it out!

It follows that

$$\int_0^{\frac{\pi}{2}} \sin^2(\theta)d\theta = \frac{1}{2}\int_0^{\frac{\pi}{2}} d\theta - \frac{1}{2}\int_0^{\frac{\pi}{2}} \cos(2\theta)d\theta$$

$$= \frac{\pi}{4} - \frac{1}{4}\int_0^{\pi} \cos(\theta)d\theta = \frac{\pi}{4}.$$

It follows that the expected value of the length in this case is

$$2\cdot\frac{\pi}{4} = \frac{\pi}{2}$$

which DOES NOT match what we got in the Possibility One part above as $\frac{\pi}{2} < \frac{2\pi}{3}$! How disturbing! Or is it?

To understand why the discrepancy between Possibility One and Possibility Two is reasonable, think about the problem from a visual point of view. In Possibility One, all lengths are equal. There is nothing in the mechanism to prefer one length over another. In Possibility Two, however, there is a hidden weight of sorts. If we fix

the point (x, y) inside the disk where the chord connecting the two endpoints of the arc and the perpendicular through the center of the circle connect, the length of the arc equals $2 \cdot \cos^{-1}(\sqrt{x^2 + y^2})$, the function which decreases with the length of (x, y). However, to compute the expected value, we integrate over the disk. The amount of area corresponding to larger values of the radius is greater. To see what precisely this means, divide D, the disk of radius 1, into annuli

$$A_k = \left\{(x, y) : k\epsilon \leq \sqrt{x^2 + y^2} \leq (k + 1)\epsilon\right\},$$

where ϵ is a very small number. Here k ranges from, say, 1 to approximately ϵ^{-1}. We have

$$\text{area } (A_k) = \pi(((k + 1)\epsilon)^2 - (k\epsilon)^2) = \epsilon^2\pi(2k + 1).$$

What this shows is that while the radius changes by the same amount inside each annulus A_k, the amount of area increases linearly with the radius of A_k. As we point out above, a LARGER radius leads to a SMALLER length of the corresponding arc, so smaller arcs are WEIGHED MORE! This explains on an intuitive level why the expected value of the length of the arc with respect to the probability function in Possibility Two is smaller than the one with respect to the probability function in Possibility One. You should keep this example in mind when reading about or studying statistics. Quite a few people make statements like: "the probability that this and that happens is so and so...". Well, it very much depends on what probability function you are using and the art and science of applied statistics is often to come up with the right function, the one that reflects the reality of the situation instead of the inevitable personal biases of the speaker.

9.6. Inclusion-Exclusion.

9.6.1. *Discrete inclusion-exclusion.* Let A and B be arbitrary finite sets. What is $\#(A \cup B)$? If $A \cap B = 0$, it is clear that $\#(AUB) = \#A + \#B$. However, if A and B intersect, the count changes. Think about it from an algorithmic point of view. Think of A as objects in a bin labeled A and B as objects in a bin labeled B. We take an item out of A and ask ourselves whether there is the same item in B. If the answer is NO, we add 1 to the column measuring the number

of elements in $A \cup B$. If the answer is YES, we still add 1 to this column, but we also take the duplicate object out of B, toss it into a trash can outside and forget about it. After we are done with A, we have $\#A$ is in the column measuring the size of $A \cup B$ and we have $\#B - \#(A \cap B)$ elements left in B. We now put 1 in our column for every one of those remaining elements and we see that we have proved that

$$\#(A \cup B) = \#A + \#B - \#(A \cap B). \tag{9.21}$$

What happens if there are more sets? Let A, B, C be sets and let us try to play the same game again. We start with A and take its elements out one at a time. If an element of A is not contained in either B or C, we simply put a 1 in our column measuring the size of $A \cup B \cup C$. If an element of A is in either B or C (or both), we put a 1 in the column and toss the duplicate elements out. After we are done with A, we have $\#A$ 1s in our column, $\#B - \#(A \cap B)$ elements left in B and $\#C - \#(A \cap C)$ elements left in C. Fantastic! But we are not done... We now work on B. If a remaining element of B is not in C, we simply put a 1 in our column. If it is in C, we put a 1 in our column and remove the duplicate from C. When we are done with B, we have $\#A + \#B - \#(A \cap B)$ elements in our column and $\#C - \#(A \cap C) - \#(B \cap C)$ elements in C. But is that really true? NO! Because if an element of $B \cap C$ is also in A, it was already removed! This means that when we are done with B, C really contains

$$\#C - \#(A \cap C) - \#(B \cap C) + \#(A \cap B \cap C)$$

elements. This leads to the conclusion that

$$\#(A \cup B \cup C) = \#A + \#B + \#C$$
$$-\#(A \cap B) - \#(A \cap C) - \#(B \cap C) + \#(A \cap B \cap C).$$

Our immediate goal, as you have probably already guessed, is a formula for$\#(A_1 \cup A_2 \cup \cdots \cup A_n)$. What do we do? We see the pattern well enough, or so it seems. It appears that we need to subtract two-fold intersections, add on three-fold intersections, subtract four-fold intersections and so on. How do we prove it? We must show that

$$\#(A_1 \cup A_2 \cup \cdots \cup A_n) \tag{9.22}$$

$$= \sum \#A_j - \sum A_{i1} \cap A_{i2} + \cdots + (-1)^{n+1} \#(A_{i1} \cap A_{i2} \cap \cdots \cap A_{in}).$$

We are going to give a very sneaky proof, but one that provides a technique that is extremely useful in probability and statistics. Let $x \in A_1 \cup A_2 \cup \cdots \cup A_n$ and suppose that x belongs to m of the sets, $1 < m < n$. Then x contributes 1 to the left hand side of the equation (9.22) and it suffices for us to show that it also contributes 1 to the right hand side. Well, x contributes m to the first sum, $-C(2, m)$ to the second sum, $C(3, m)$ to the third sum, ..., $(-1)^{k+1}$ to the kth sum, $k < m$, and does not contribute anything to the remaining sums. Thus the contribution of x to the right hand side is

(9.23)

$$C(1, m) - C(2, m) + \cdots + (-1)^{m+1} C(m, m) = \sum_{k=1}^{m} (-1)^{k+1} C(k, m).$$

By the binomial formula (Theorem 9.3 above),

$$0 = (1-1)^m = \sum_{k=0}^{m} (-1)^k C(k, m) = \sum_{k=0}^{m} (-1)^{k+1} C(k, m)$$

$$= -1 + \sum_{k=1}^{m} (-1)^{k+1} C(k, m),$$

and we conclude that

$$\sum_{k=1}^{m} (-1)^{k+1} C(k, m) = 1.$$

In view of (9.23) and the preceding discussion, we conclude that x contributes exactly 1 to the right hand side of (9.22) and the proof is complete.

9.6.2. *Continuous inclusion-exclusion.* There are many things we can do- a constant problem in mathematics- but we must choose something and restrain ourselves. Let $A_1, A_2, \ldots, A_n$ be a collection of open Riemann measurable subsets of the plane. Recall that this simply means that we can integrate over them... Also recall that the set is open means that if a point is in the set, then a sufficiently small ball centered at this point is also contained in the set. Let $|A_j|$ denote

the area of A_j, or, more precisely

$$\int\int_A 1 dxdy.$$

In analogy with the discrete case, we would like to compute

$$|A_1 \cup A_2 \cup \cdots \cup A_n|,$$

and every instinct we have tells us that the answer is the same,

$$|A_1 \cup A_2 \cup \cdots \cup A_n|$$

$$= \sum |A_j| - \sum |A_{i_1} \cap A_{i_2}| + \cdots + (-1)^{n+1} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_n}|. \tag{9.24}$$

The counting argument we gave above no longer works. Or does it? We are assuming that each A_j is open, so if $(x, y) \in A_j$, there exists an $\epsilon > 0$ such that the disk of radius ϵ centered at (x, y) is also contained in A_j. Suppose that (x, y) belongs to m of the sets, with $1 < m < n$ and choose ϵ small enough so that the ϵ disk centered at (x, y) is contained in all of the sets that (x, y) is contained in. Now, this disk contributes $\pi\epsilon^2$ to the left hand of (9.24). We must prove that it contributes the same amount to the right hand side. But it does! The proof is exactly the same. Instead of contributing m to the first sum on the right hand side, the disk contributes in $m \cdot \pi\epsilon^2$. Instead of contributing $-C(2, m)$ to the second sum on the right hand side, the disk contributes $-C(2, m) \cdot \pi\epsilon^2$ and so on. The proof is really the same! This raises the question of whether there is a more general way to formulate everything we are doing in order to combine discrete and continuous situations and of course there is. One of the purposes of the way things are structured here is to make you yearn for a more general theory and enjoy it immensely when it finally arrives.

I have left out applications of inclusion-exclusion for the time being and I did it on purpose! Can you come up with some simple examples illustrating the immense usefulness of this formula? The answer is, YES, you can. So just do it!

1. Notes, remarks and difficult questions

In order to really get a feel for continuous probability, it pays to workout several computationally intensive examples. Here is one that

quickly leads to an open problem. Let K be a convex planar set. Let $p_K(r)$ denote the probability that a needle of length r whose center has equal probability of landing anywhere in K is completely contained in K.

Exercise 9.1. Compute $p_K(1)$ when K is a square of side-length one and also in the case when K is the disk of radius $\frac{1}{\sqrt{\pi}}$. Which one is larger?

This leads us to a pretty unsolved problem in integral geometry.

Exercise 9.2. Prove that among all the convex sets of area one, $p_K(r)$ is maximized when K is the disk, provided that r is sufficiently small.

What happens when r is large? Well, if r is large enough, $p_K(r) = 0$. In between, I would guess the maximizing shape is an ellipse of some sort, but in reality this is anybody's guess at this point.

Chapter 10

A more involved taste of probability

We have explored discrete, continuous and arithmetic mathematics in various parts of this book and even made some attempts to combine these notions. We continue this process in this chapter where we attempt to understand the following basic question. Suppose that two positive integers are picked at "random". How likely is it that they are relatively prime? Recall that two integers are relatively prime if they do not have any common factors. The probability that a given integer is divisible by a prime p is $\frac{1}{p}$. The probability that two integers are both divisible by p is $\frac{1}{p^2}$. The probability that at least one of them is not divisible by p is $1 - \frac{1}{p^2}$.

It follows that the probability that two positive integers are relatively prime is

$$\Pi_{p \in \mathcal{P}} \left(1 - \frac{1}{p^2}\right), \tag{10.1}$$

where $\mathcal{P}$ is the set of positive primes. DO NOT just accept this! Why are we multiplying? Why? Answer this question to your complete satisfaction or keep thinking about it! Just what do I mean by probability here? How do I make this rigorous? A small hint... Instead of considering all positive integers, consider ones between 1 and N, N

very large. Now it is more or less clear what we mean by probability and we then take a limit as $N \to \infty$.

What is this mysterious quantity in (10.1)? We can rewrite it in the form

$$\frac{1}{\Pi_{p\in\mathcal{P}} \frac{1}{1-p^{-2}}}. \tag{10.2}$$

Now,

$$\frac{1}{1-p^{-2}} = 1 + p^{-2} + p^{-4} + \dots, \tag{10.3}$$

is the geometric series formula. To see this, let

$$s = 1 + p^{-2} + p^{-4} + \dots.$$

Then

$$p^{-2}s = p^{-2} + p^{-4} + \dots.$$

Subtracting the two gives us

$$s(1 - p^{-2}) = 1,$$

and the claim follows.

Plugging (10.3) into (10.2) we get

$$\frac{1}{\Pi_{p\in\mathcal{P}} \sum_{k=0}^{\infty} p^{-2k}}.$$

Now, the expression is still goofy looking. Can we make it look a bit more humane? What is

$$\Pi_{p\in\mathcal{P}} \sum_{k=0}^{\infty} p^{-2k}?$$

It is none other than the sum of all possible prime factorizations raised to the power of -2! This means that

$$\frac{1}{\Pi_{p\in\mathcal{P}} \sum_{k=0}^{\infty} p^{-2k}} = \frac{1}{\sum_{n=1}^{\infty} \frac{1}{n^2}}, \tag{10.4}$$

and we have an answer of sorts.

Exercise 10.1. In order to get a better feel for the formula (10.4) first consider the case when the left hand side is summed over only one prime. What happens in this case? Then try the case of two primes, and go from there.

Let us give the sum on the right hand side of (10.4) a name. It will seem completely out of context, but that is fine, believe me... For a complex number s, define

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \tag{10.5}$$

With this notation, we have just seen above that the probability that two positive integers are relatively prime is $\frac{1}{\zeta(2)}$. We mention in passing that this $\zeta(s)$ is the Riemann zeta function, one of the most difficult and important objects in mathematics. Having made this cryptic remark, we return to the business at hand. In order to make sense of what we have done, we need to have an idea of just what sort of a number this $\zeta(2)$ is. As usual, we are going to play around a bit, a habit we are trying very hard to encourage!

Observe that $n^2 > n(n-1)$. It follows that $\frac{1}{n^2} < \frac{1}{n(n-1)}$. We conclude that

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{n^2} &< 1 + \sum_{n=2}^{\infty} \frac{1}{n(n-1)} \\ &= 1 + \sum_{n=2}^{\infty} \left(\frac{1}{n-1} - \frac{1}{n} \right) \\ = 1 + \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \ldots 1 + 1 &= 2. \end{aligned} \tag{10.6}$$

It follows that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \leq 2,$$

and the conclusion, in light of our work above, is that the probability that two positive integers are relatively prime is $\geq \frac{1}{2}$. This is already quite remarkable. We are saying that if we pick two positive integers, there is a better than 50-50 chance that they do not have any factors

in common. It turns out, however, that the probability is bigger than that!

Let us continue exploring. We have

$$\sum_{n=1}^{\infty} \frac{1}{n^2}$$

$$= 1 + \frac{1}{4} + \sum_{n=3}^{\infty} \frac{1}{n^2} \leq \frac{5}{4} + \sum_{n=3}^{\infty} \frac{1}{n(n-1)}$$

$$= \frac{5}{4} + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \dots$$

$$\frac{5}{4} + \frac{1}{2} = \frac{7}{4} < 2 \dots$$

We can do even better... We have

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{4} + \frac{1}{9} + \sum_{n=4}^{\infty} \frac{1}{n^2}$$

$$\leq \frac{11}{8} + \sum_{n=4}^{\infty} \frac{1}{n(n-1)}$$

$$= \frac{11}{8} + \left(\frac{1}{3} - \frac{1}{4}\right) + \left(\frac{1}{4} - \frac{1}{5}\right) + \dots$$

$$\frac{11}{8} + \frac{1}{3} = \frac{41}{24} < \frac{7}{4} \dots$$

We can keep improving this bound, but we do not have all day, so to speak! So what is the answer?! Let us go for it... The argument we are about to present is due to Euler and I learned it from Leonardo (Maestro) Colzani's notes (in Italian) on the history of mathematics.

By the chain rule, the derivative of $(\arcsin(x))^2$ is $\frac{2\arcsin(x)}{\sqrt{1-x^2}}$, from which it follows by the Fundamental Theorem of Calculus that

$$\int_0^1 \frac{\arcsin(x)}{\sqrt{1-x^2}} dx \tag{10.7}$$

$$= \frac{1}{2}((\arcsin(1))^2 - (\arcsin^2(0))^2) = \frac{\pi^2}{8}.$$

Notice that we have used the fact that the derivative of $\arcsin(x)$ is $\frac{1}{\sqrt{1-x^2}}$. Let us recall how this goes. If $y = \arcsin(x)$, then $x = \sin(y)$. Differentiating both sides with respect to x, using the chain

rule, we see that $1 = \cos(y) \cdot y'$. It follows that $y' = \frac{1}{\cos(y)}$. What is $\cos(y)$? Well, $x = \sin(y)$ and $\sin^2(y) + \cos^2(y) = 1$, so $\cos(y) = \sqrt{1 - x^2}$. We conclude that

$$(\arcsin(x))' = \frac{1}{\sqrt{1 - x^2}},$$

as claimed.

Let us now explore further and get a Taylor expansion at the origin for $\arcsin(x)$. Yes, I know, this was a long time ago. We shall do it in a way that causes us to learn something mildly amusing, as usual. The Binomial Theorem tells us how to expand $(a + b)^n$ when n is a positive integer. Isaac Newton worked out a variant of the Binomial Theorem for $(1 + x)^a$, where x is any real number and a is any rational, positive or not! He showed that

$$(1 + x)^a = 1 + ax + \frac{a(a-1)}{2!}x^2 + \frac{a(a-1)(a-2)}{3!}x^3 + \dots. \tag{10.8}$$

This is a direct corollary of Taylor's theorem, which in the case when things are centered at the origin says that if the function f is infinitely differentiable, then

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2!}f''(0) + \dots + \frac{x^n}{n!}f^{(n)}(0) + \dots.$$

Replace x by $-x^2$ and set $a = -\frac{1}{2}$. Plugging this into (10.8) we get

$$(1 - x^2)^{-\frac{1}{2}} = 1 + \frac{1}{2}x^2 + \frac{1}{2!}\frac{1 \cdot 3}{2 \cdot 2}x^4 + \frac{1}{3!}\frac{1 \cdot 3 \cdot 5}{2 \cdot 2 \cdot 2}x^6 + \dots.$$

Integrating this expression we see that

$$\arcsin(x) = x + x^3\frac{1}{2 \cdot 3} + x^5\frac{1}{2!}\frac{1 \cdot 3}{2^2 \cdot 5} + x^7\frac{1}{3!}\frac{1 \cdot 3 \cdot 5}{2^3 \cdot 7} + \dots.$$

Putting our calculations back into (10.7) we see that

$$\frac{\pi^2}{8} = \int_0^1 \frac{\arcsin(x)}{\sqrt{1 - x^2}}dx \tag{10.9}$$

$$= \int_0^1 \frac{x}{\sqrt{1 - x^2}} + \frac{1}{2 \cdot 3}\int_0^1 \frac{x^3}{\sqrt{1 - x^2}}dx + \frac{1}{2!}\frac{1 \cdot 3}{2^2 \cdot 5}\int_0^1 \frac{x^5}{\sqrt{1 - x^2}}dx + \dots$$

which brings us to the highly appetizing prospect of computing

$$\int_0^1 \frac{x^{2k+1}}{\sqrt{1-x^2}}dx, \tag{10.10}$$

for $k = 0, 1, 2 \ldots$ It is a bitch, but it needs to be done. Let $x = \sin(\theta)$. Then $dx = \cos(\theta)d\theta$. We see that the expression in (10.10) equals

$$N_k = \int_0^{\frac{\pi}{2}} \sin^{2k+1}(\theta)d\theta = \int_0^{\frac{\pi}{2}} \sin^{2k}(\theta)\sin(\theta)d\theta. \tag{10.11}$$

We now integrate by parts. Let $u = \sin^{2k}(\theta)$and $dv = \sin(\theta)d\theta$. Recalling the integration by parts formula

$$\int udv = uv - \int vdu,$$

we see that the expression in (10.11) equals

$$-\sin^{2k}(\theta)\cos(\theta)\mid_0^{\frac{\pi}{2}} + 2k\int_0^{\frac{\pi}{2}} \sin^{2k-1}(\theta)\cos^2(\theta)d\theta$$

$$= 2k\int_0^{\frac{\pi}{2}} \sin^{2k-1}(\theta)(1-\sin^2(\theta))d\theta$$

$$= 2k\int_0^{\frac{\pi}{2}} \sin^{2k-1}(\theta)d\theta - 2k\int_0^{\frac{\pi}{2}} \sin^{2k+1}(\theta)d\theta$$

$$= 2k(N_{k-1} - N_k).$$

Thus we have derived a cute looking recursion

$$N_k = 2k(N_{k-1} - N_k),$$

or, equivalently,

$$\frac{N_k}{N_{k-1}} = \frac{2k}{2k+1}. \tag{10.12}$$

What do we do with this? Well, observe that we have a telescoping product on our hands! Where, you ask... Well,

$$\Pi_{k=1}^m \frac{N_k}{N_{k-1}} = \frac{N_1}{N_0}\cdot\frac{N_2}{N_1}\cdot\frac{N_3}{N_2}\cdots\frac{N_m}{N_{m-1}} = \frac{N_m}{N_0} = N_m$$

since we can verify by a direct calculation (do it!) that $N_0 = 1$. It now follows from (10.12) that for $m \geq 1$,

$$N_m = \Pi_{k=1}^m \frac{2k}{2k+1} = \frac{2}{3}\cdot\frac{4}{5}\cdots\frac{2m}{2m+1}. \tag{10.13}$$

It follows that we get

$$N_0 + \sum_{m=1}^{\infty} \frac{1}{m!} \frac{1 \cdot 3 \ldots 2m-1}{2^m \cdot (2m+1)} N_m.$$

Plugging in (10.13) we get

$$N_0 + \sum_{m=1}^{\infty} \frac{1}{m!} \frac{1 \cdot 3 \ldots 2m-1}{2^m \cdot (2m+1)} \Pi_{k=1}^{m} \frac{2k}{2k+1}$$

$$= N_0 + \sum_{m=1}^{\infty} \frac{1}{m!} \frac{1 \cdot 3 \ldots 2m-1}{2^m \cdot (2m+1)} \cdot \frac{2}{3} \cdot \frac{4}{5} \cdots \frac{2m}{2m+1}$$

$$= 1 + \sum_{m=1}^{\infty} \frac{1}{(2m+1)^2} = \sum_{m=0}^{\infty} \frac{1}{(2m+1)^2},$$

since, as we asked you to verify above, $N_0 = 1$.

Thus we have derived the formula

$$\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \ldots.$$

How do we get from this formula to the one for

$$\mathrm{X} = \sum_{n=1}^{\infty} \frac{1}{n^2}?$$

Well,

$$\mathrm{X} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \ldots$$

$$= \left(1 + \frac{1}{3^2} + \frac{1}{5^2} + \ldots\right) + \frac{1}{4}\left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \ldots\right)$$

$$= \frac{\pi^2}{8} + \frac{\mathrm{X}}{4}.$$

It follows that

$$\mathrm{X} = \frac{\pi^2}{8} + \frac{\mathrm{X}}{4},$$

which means that

$$\frac{3 \cdot \mathrm{X}}{4} = \frac{\pi^2}{8},$$

so

$$\mathrm{X} = \frac{\pi^2}{6}.$$

Finally!!! At last we know that the probability that two positive integers are relatively prime is

$$\frac{1}{\frac{\pi^2}{6}} = \frac{6}{\pi^2},$$

which is close to $\frac{2}{3}$.

As always, there is more to do. What is the probability that three randomly chosen positive integers are relatively prime, four, five, ..., k randomly chosen positive integers? It is not hard to modify the argument above (do it!) to see that the probability that k randomly chosen positive integers are relatively prime is $\frac{1}{\zeta(k)}$, where ζ is defined above.

What is much harder to do is to actually evaluate $\zeta(k)$ for any k other than 2. It turns out that if k is even, $\zeta(k)$ can be evaluated, but it is very difficult. Try it for $k = 4$ if you really enjoy pain. For k odd, it is not even known in some cases if $\zeta(k)$ is rational or not. The fact that $\zeta(3)$ is irrational was only proved in the late 80s and it is still not known whether or not it is transcendental!

1. Notes, remarks and difficult questions

In addition to establishing a pretty probabilistic estimate, this chapter introduces the Riemann zeta function, one of the key objects of modern mathematics. The Riemann Hypothesis, one of the greatest unsolved problems for which the Clay Institute is offering a million dollar prize, says that

$$\zeta(s) = 0 \text{ if and only if } Re(s) = \frac{1}{2}.$$

In this chapter we have seen that properties of the Riemann zeta function are intimately connected with the properties of prime numbers. It turns out that the Riemann Hypothesis is equivalent to the statement that if P is the set of all positive prime numbers, then

$$\pi(x) = \#\{n \in P : n \leq x\} = \frac{x}{\log(x)} + O(x^{\frac{1}{2}+\epsilon})$$

for any $\epsilon > 0$.

A much weaker but already an extremely deep and important fact, that

$$\frac{\pi(x)}{\frac{x}{\log(x)}} \to 1 \text{ as } x \to \infty$$

is known as the Prime Number Theorem, proved by Hadamard and, independently, by Poussin, at the end of the 19th century. See, for example, [**1**] and the references contained therein.

What is somewhat more bizarre is that the Riemann Hypothesis turns out to be equivalent to the following strange looking statement. We say that the height of a rational number t in $[0,1]$, $h(t)$, written in reduced form, is equal to its denominator. Then the Riemann Hypothesis is equivalent to the estimate

$$\left| \frac{1}{H^{\frac{3}{2}+\epsilon}} \sum_{h(t)\leq H} e^{2\pi i t} \right| \to 0 \text{ as } H \to \infty.$$

It is also amusing to note that proving a weaker fact that

$$\left| \frac{1}{H^2} \sum_{h(t)\leq H} e^{2\pi i t} \right| \to 0 \text{ as } H \to \infty$$

is already equivalent to the Prime Number Theorem mentioned above. See, for example, [**8**], and the references contained therein for a more detailed description and applications.

Chapter 11

Oscillatory integrals and fun that lies beyond

Let me say straight out that the significance of this chapter will not be clear until the next one... Nevertheless, I hope that you enjoy the way the mean value theorem allows us to accurately estimate seemingly intractable integrals, and how complex integration connects new integrals we encounter here with some of our old friends from the preceding chapters. Nothing beyond first, and a tiny of bit of second, year calculus is required here, except for one small part... where complex integration is used. I also wish to note in passing that I was inspired to write this chapter after taking another look at one of the best calculus books ever written, the one by Edmund Landau. In that book he presents some of the material of this chapter as an application of the power of the mean value theorem in sharp contrast to the great treatises of today where applications are frequently comical at best.

11.1. Basic oscillatory integrals. Recall that the mean value theorem says that if $f(x)$ is differentiable on (a, b) and continuous on $[a, b]$, then there exists $c \in (a, b)$ such that

$$f(b) - f(a) = (b - a)f'(c).$$

Another familiar theme from first year calculus is that one looks for maxima and minima of functions by examining the first derivative

$f'(x)$. In the chapter we use these basic principles to study integrals of the form

$$I_f(R) = \int_a^b e^{iRf(x)} dx,$$

where f is a suitably differentiable function and R is a large parameter. More precisely, we will see that for a reasonable class of functions f,

$$I_f(R) \to 0 \text{ as } R \to \infty,$$

and in many cases we shall determine the rate of this convergence. What is the idea here? Since

$$e^{ix} = \cos(x) + i\sin(x), \tag{11.1}$$

$$|e^{ix}| = \sqrt{\cos^2(x) + \sin^2(x)} = 1.$$

It follows that

$$\left| \int_a^b e^{iRf(x)} dx \right| \leq \int_a^b |e^{ifR(x)}| dx = \int_a^b dx = (b-a). \tag{11.2}$$

The question is, can we do better? Suppose that $f(x) = x$. Then

$$I_f(R) = \int_a^b e^{iRx} dx = \frac{e^{ix}}{iR}|_a^b$$

$$= \frac{e^{ib} - e^{ia}}{iR},$$

and the absolute value of this quantity is $\leq \frac{1}{R}$, which beats (12.13) the moment $R > \frac{1}{b-a}$.

We would like to be able to extend this type of an estimate to a wider class of functions. To do that, we need to understand which properties of the function $f(x) = x$ we are really using. We cannot rely on the possibility of integrating $e^{iRf(x)}$ explicitly, because, as we may or may not have been told when we took calculus, being able to integrate a function in closed form is a man bites dog situation! It seldom happens, except in contrived problems designed for content challenged college mathematics texts. However, even those texts give us a hint of what to do with functions that are difficult to integrate.

After pedestrian trickery like u-substitution fails, we typically turn to integration by parts. Let us give it a try here. Observe that

$$\frac{d}{dx}\left(e^{iRf(x)}\right) = iRf'(x)e^{iRf(x)}.$$

It follows that

$$I_f(R) = \int_a^b \frac{1}{iRf'(x)} \frac{d}{dx}\left(e^{iRf(x)}\right) dx.$$

Integrating by parts we see that this quantity equals

$$\frac{e^{iRf(x)}}{iRf'(x)}|_a^b - \int_a^b e^{iRf(x)} \frac{d}{dx}\left(\frac{1}{iRf'(x)}\right) dx = I + II.$$

Suppose that we knew that

$$f'(x) \geq 1. \tag{11.3}$$

We could then conclude that

$$|I| = \left|\frac{e^{iRf(b)}}{iRf'(b)} - \frac{e^{iRf(a)}}{iRf'(a)}\right| \leq \frac{2}{R}.$$

This is very encouraging as it is beginning to look like we may be able to get a $\frac{1}{R}$ type bound by assuming (11.3). Let us keep trucking along and see. We could differentiate the expression inside II and see what happens, but that looks highly unappealing... Let us get greedy instead and just take absolute values inside. More precisely,

$$|II| \leq \int_a^b \left|\frac{d}{dx}\left(\frac{1}{iRf'(x)}\right)\right| dx.$$

Before we get upset at this seemingly intractable monster, let us fantasize a little bit. Suppose that absolute value just disappeared like thieves after a successful robbery. We could then use the fundamental theorem of calculus (do it!) and obtain the same estimate as we did for I above. How do we make absolute values disappear? Well, if the function under the absolute values were ≥ 0 or strictly ≤ 0, we would need absolute values like a hole in the head. We are on a roll, so let us strain our mental faculties further. How do we know if a quantity being differentiated is strictly ≥ 0 or ≤ 0? Again, this is not rocket science. If the quantity is strictly increasing or strictly decreasing, its derivative is strictly positive or negative, respectively.

It follows that if we assume that f' is strictly increasing or strictly decreasing, then

$$|II| \leq \int_a^b \left| \frac{d}{dx}\left(\frac{1}{iRf'(x)}\right)\right| dx$$
$$= \left| \int_a^b \frac{d}{dx}\left(\frac{1}{iRf'(x)}\right) dx \right| = \left| \frac{1}{iRf'(x)}|_a^b \right| \leq \frac{2}{R},$$

if we also assume (11.3). Thus we have arrived at the following result proved almost a century ago by van der Corput.

Theorem 11.4. *Let f be a once differentiable function such that f' is either strictly increasing or strictly decreasing and (11.3) holds. Then*

$$|I_f(R)| \leq \frac{4}{R}.$$

Exercise 11.1. Does the conclusion of Theorem (11.4) still hold if (11.3) is removed? What about the monotonicity assumption? We shall address the former in a moment, but you should definitely think about the latter.

11.2. Justifying the assumptions. We now address the necessity of (11.3) assumption in earnest in order to justify the rest of this chapter. How do we show that (11.3) is necessary? One way is to exhibit a differentiable function f for which (11.3) DOES NOT hold and the conclusion of Theorem 11.4 DOES NOT hold either. This is our plan!

Let $f(x) = x^2$, $a = 0$ and $b = 1$. We have

$$I_f(R) = \int_0^1 e^{iRx^2} dx = \frac{1}{\sqrt{R}} \int_0^{\sqrt{R}} e^{ix^2} dx$$
$$= \frac{1}{\sqrt{R}} \left(\int_0^\infty e^{ix^2} dx - \int_{\sqrt{R}}^\infty e^{ix^2} dx \right).$$

Suppose that we could show that

$$\left| \int_0^\infty e^{ix^2} dx \right| \geq C_1 > 0 \tag{11.5}$$

and

$$\left| \int_{\sqrt{R}}^\infty e^{ix^2} dx \right| \leq \frac{C}{\sqrt{R}}. \tag{11.6}$$

We would then be forced to conclude, (why?) in particular, that for any $\epsilon > 0$,

$$R^{\frac{1}{2}+\epsilon}|I_f(R)| \to \infty \text{ as } R \to \infty,$$

so the best estimate we can hope, and indeed obtain, in this situation, is

$$|I_f(R)| \leq \frac{C}{\sqrt{R}}$$

for some positive constant R. In particular, the conclusion of Theorem 11.4 is just not in the cards... Do you see what we are up to here? We obtained an estimate on $I_f(R)$ based on a set of reasonable assumptions on f. We then found an example, provided that we can verify (11.5) and (11.6) that shows that we cannot in general have the estimate as good as (11.4). What do you think we are going to do next? Stay tuned... and help me verify (11.5) and (11.6).

We first take care of (11.6). Write

$$\int_{\sqrt{R}}^{\infty} e^{ix^2} dx$$

$$= \int_{\sqrt{R}}^{\infty} \frac{d}{dx}\left(e^{ix^2}\right) \frac{dx}{2ix}$$

$$= \frac{e^{ix^2}}{2ix}|_{\sqrt{R}}^{\infty} + \frac{1}{2i}\int_{\sqrt{R}}^{\infty} e^{ix^2}\frac{dx}{x^2} = I + II.$$

Using the fact that $|e^{i\ whatever}| = 1$,

$$|I| \leq \frac{1}{2\sqrt{R}}.$$

Exercise 11.2. Recall that if $z = a + ib$ is a complex number, then $|z| = \sqrt{a^2 + b^2}$ and this is what we mean by absolute values above. Go over everything with this in mind and get it together...

Now,

$$|II| \leq \frac{1}{2}\int_{\sqrt{R}}^{\infty} \left| e^{ix^2}\frac{1}{x^2}\right| dx$$

$$\leq \frac{1}{2}\int_{\sqrt{R}}^{\infty} \frac{dx}{x^2} \leq \frac{1}{2\sqrt{R}}.$$

We conclude that

$$\left|\int_{\sqrt{R}}^{\infty} e^{ix^2}dx\right| \leq \frac{1}{\sqrt{R}},$$

thus establishing (11.6).

We must now prove (11.5). We will do this in two different ways. The first method, which will actually yield the exact value of the integral, uses complex analysis. The second method is completely elementary, though it does not yield the exact value of the integral, and is outlined in the Exercise 11.4 below.

Let $\gamma = \gamma_1 \cup \gamma_2 \cup \gamma_3$, where

$$\gamma_1 = \{z = t; 0 \leq t \leq R\},$$

$$\gamma_2 = \left\{z = Re^{i\theta}; 0 \leq \theta \leq \frac{\pi}{4}\right\},$$

and

$$\gamma_3 = \left\{z = \frac{t+it}{\sqrt{2}}; 0 \leq t \leq R\right\}.$$

Since the function $f(z) = e^{iz^2}$ is analytic inside γ, and γ is a closed contour, we have, by elementary complex analysis,

$$\int_{\gamma} e^{iz^2}dz = 0.$$

On the other hand,

$$\int_{\gamma} e^{iz^2}dz = \int_{\gamma_1} e^{iz^2}dz + \int_{\gamma_2} e^{iz^2}dz + \int_{\gamma_3} e^{iz^2}dz.$$

Plugging in the definition, we see that

$$\int_{\gamma_1} e^{iz^2}dz = \int_0^R e^{it^2}dt.$$

This is excellent because this is the integral that we want once we take $R \to \infty$. Plugging in the second definition, we see that

$$\left|\int_{\gamma_2} e^{iz^2}dz\right| = \left|\int_0^{\frac{\pi}{4}} e^{iR^2e^{2i\theta}} Re^{i\theta}d\theta\right|$$

$$\left|\int_0^{\frac{\pi}{4}} e^{iR^2(\cos(2\theta)+i\sin(2\theta))} Re^{i\theta}d\theta\right|$$

$$\leq R\int_0^{\frac{\pi}{4}} e^{-R^2\sin(2\theta)}d\theta$$

$$= R\int_0^{\delta} e^{-R^2\sin(2\theta)}d\theta + R\int_{\delta}^{\frac{\pi}{4}} e^{-R^2\sin(2\theta)}d\theta = I + II.$$

Now,

$$I \leq R\delta,$$

while

$$II \leq \frac{\pi}{4}\cdot R\cdot e^{-R^2\sin(2\delta)}$$

since $\sin(2\theta)$ is an increasing function on the interval $[0, \frac{\pi}{4}]$.

Choosing $\delta = R^{-\frac{3}{2}}$, for example, makes both I and II tend to 0 as $R \to \infty$.

Exercise 11.3. Carefully verify all the steps in the estimation of $\int_{\gamma_2} e^{iz^2}dz$ above. For example, why did we choose δ the way we did? Could we have chosen a power other than $-\frac{3}{2}$. If so, what other powers could we have chosen?

Having driven the second integral to 0, in the limit as $R \to \infty$, we are nearing the moment of truth. Plugging in the third definition, we have

$$-\int_{\gamma_3} e^{iz^2}dz = -\frac{1+i}{\sqrt{2}}\cdot\int_0^R e^{\frac{i(t+it)^2}{2}}dt$$

$$= -\frac{1+i}{\sqrt{2}}\int_0^R e^{-t^2}dt.$$

In the limit as $R \to \infty$ we obtain

$$-\frac{1+i}{\sqrt{2}}\int_0^{\infty} e^{-t^2}dt$$

$$= -\frac{1+i}{2\sqrt{2}}\int_{-\infty}^{\infty} e^{-t^2}dt$$

$$= -\frac{\sqrt{\pi}(1+i)}{2\sqrt{2}}$$

by the calculation for I_1, see (4.3), we performed in Chapter 4!

We conclude that

$$\int_0^{\infty} e^{it^2}dt = \frac{\sqrt{\pi}(1+i)}{2\sqrt{2}}. \tag{11.7}$$

This verifies (11.5). We note in passing that (11.7) implies that

$$\int_0^\infty \cos(x^2)dx = \frac{\sqrt{\pi}}{2\sqrt{2}},$$

and

$$\int_0^\infty \sin(x^2)dx = \frac{\sqrt{\pi}}{2\sqrt{2}}.$$

We now outline an elementary approach to establishing (11.5) though it does not yield the exact value of the integral.

Exercise 11.4. Write

$$\int_0^\infty \cos(x^2)dx = \int_0^\delta \cos(x^2)dx + \int_\delta^\infty \cos(x^2)dx.$$

Convince yourself that it is enough to deal with $\cos(x^2)$ instead of e^{ix^2}. Then estimate the first integral from below, using the fact that $\cos(x^2)$ is positive near 0 and the second integral from above using integration by parts. If you choose δ appropriately, you will succeed. Please allocate some time for this exercise...

11.3. Second derivative makes an entrance. We are now convinced that if the first derivative of f is not bounded from below, $I_f(R)$ cannot behave as well as Theorem 11.4 prescribes. This only begs the question... What happens if the second derivative is bounded from below. Let's try it and see! If $f''(x) \geq 1$, we may still get lucky and actually have that $f'(x) \geq 1$. In such a case we simply apply Theorem 11.4. What happens if we are not so lucky? Since $f''(x) \geq 1$, $f'(x)$ can vanish at at most one point. Theorem 11.4 suggests that we should avoid this point like a plague! How far should one get away from the plague? That all depends on the equipment you carry... Let's explore! We have

$$\int_a^b e^{iRf(x)}dx = \int_a^{plague-\delta} e^{iRf(x)}dx$$

$$+ \int_{plague-\delta}^{plague+\delta} e^{iRf(x)}dx + \int_{plague+\delta}^b e^{iRf(x)}dx = A + B + C,$$

where the "sanitary" distance δ will be determined in due time.

Common sense tells us that we will need to be brave and get pretty close to the plague. In other words, δ will have to be pretty

small, so the best we can do with B is to estimate it head on. We have

$$|B| \leq \int_{plague-\delta}^{plague+\delta} dx = 2\delta.$$

How do we deal with A and C. They are basically the same (Exercise 11.5 below...), so let us deal with A. By the mean value theorem,

$$f'(x) = f'(x) - f'(plague) = (x - plague)f''(c),$$

where $c \in (x, plague)$. Since $f''(c) \geq 1$, we conclude that

$$f'(x) \geq \delta$$

on $(a, plague - \delta)$.

What can we do with that? Well, Theorem 11.4 tells us what to do if $f'(x) \geq 1$ and f' is monotonic. Now, f' is monotonic alright since $f''(x) \geq 1$, but what about the problem of 1 being replaced by δ? We get around this difficulty by a simple content-free dodge. Observe that if $f'(x) \geq \delta$, then if we define

$$g(x) = \frac{f(x)}{\delta},$$

$$g'(x) \geq 1.$$

It follows that

$$\left| \int_a^{plague-\delta} e^{iRf(x)} dx \right|$$

$$= \left| \int_a^{plague-\delta} e^{iR\delta g(x)} dx \right|$$

$$\leq \frac{4}{R\delta},$$

by Theorem 11.4!!

We conclude that

$$|A + B + C| \leq \frac{8}{R\delta} + 2\delta.$$

If we set $\delta = \frac{1}{\sqrt{R}}$, we see that this quantity is $\leq \frac{10}{\sqrt{R}}$.

We just proved the following variant of Theorem 11.4.

Theorem 11.8. *Suppose that $f''(x) \geq 1$. Then*

$$|I_f(R)| \leq \frac{10}{\sqrt{R}}.$$

Exercise 11.5. Repeat the argument we used for term A to see that it indeed applies to term C.

11.4. Exponentials dancing on the unit disk. Let D denote the unit disk in the plane and define $\chi_D(x) = 1$ if $x \in D$ and 0 if $x \notin D$. Let us define a mysterious object

$$\widehat{\chi}_D(\xi) = \int_D e^{-2\pi i x \cdot \xi} dx. \tag{11.9}$$

Exercise 11.6. Prove that $\widehat{\chi}_D(\xi)$ is radial in the sense that

$$\widehat{\chi}_D(\xi) = \widehat{\chi}_D(\mu)$$

whenever $|\xi| = |\mu|$. (Recall that $|\xi| = \sqrt{\xi_1^2 + \xi_2^2}$).

Hint: Let $\xi = R(\cos(\phi), \sin(\phi))$ and $x = t(\cos(\theta), \sin(\theta))$. Then

$$x \cdot \xi = Rt\cos(\theta - \phi)...$$

by angle addition formulae. Don't you dare take my word for this. Check everything!

We are about to see that the games we play above have some bearing on this object... Recall that Green's theorem says that if g_1 and g_2 are continuously differentiable functions on D, then

$$\int_{\partial D} g_1(x)dx_1 + g_2(x)dx_2 = \int_D \left(\frac{\partial g_2}{\partial x_1} - \frac{\partial g_1}{\partial x_2} \right) dx. \tag{11.10}$$

In light of Exercise 11.6 it is enough to take $\xi = (0, R)$ in (11.9) and consider

$$F_D(R) = \int_D e^{-2\pi i R x_2} dx_1 dx_2.$$

Let $g_2(x) \equiv 0$ and let $g_1(x) = \frac{e^{-2\pi i R x_2}}{2\pi i R}$. Applying (11.10) we see that

$$F_D(R) = -\frac{1}{2\pi i R} \int_0^{2\pi} e^{-2\pi i R \sin(\theta)} \sin(\theta) d\theta = I + II,$$

where I is the integral from 0 to π and II is the remaining integral. We handle I, as II is estimated the same way.

Let $I = I_{great} + I_{alsogreat} + I_{OK}$, where

$$I_{great} = -\frac{1}{2\pi iR}\int_0^{\frac{\pi}{4}} e^{-2\pi iR\sin(\theta)}\sin(\theta)d\theta,$$

$$I_{alsogreat} = -\frac{1}{2\pi iR}\int_{\frac{3\pi}{4}}^{\pi} e^{-2\pi iR\sin(\theta)}\sin(\theta)d\theta,$$

and

$$I_{OK} = -\frac{1}{2\pi iR}\int_{\frac{\pi}{4}}^{\frac{3\pi}{4}} e^{-2\pi iR\sin(\theta)}\sin(\theta)d\theta.$$

Let $f(\theta) = 2\pi\sin(\theta)$. For $\theta \in [0, \pi/4]$ and $\theta \in [3\pi/4, 2\pi]$, $f'(\theta)$ is monotonic and $|f'(\theta)| \geq 1$ (better, actually...). By Theorem 11.4 this implies that

$$|I_{great}| \leq \frac{4}{2\pi R^2}, \text{ and } |I_{alsogreat}| \leq \frac{4}{2\pi R^2}.$$

On the other hand, for $\theta \in [\pi/4, 3\pi/4]$, $|f''(\theta)| \geq 1$, so by Theorem 11.8,

$$|I_{OK}| \leq \frac{10}{2\pi R^{\frac{3}{2}}}.$$

Putting everything together we see that there is a positive constant C such that

$$|F_D(R)| \leq \frac{C}{R^{\frac{3}{2}}},$$

or, in other words, we have the following theorem.

Theorem 11.11. *Let D, χ_D and $\widehat{\chi}_D$ be defined as above. Then*

$$|\widehat{\chi}_D(\xi)| \leq C|\xi|^{-\frac{3}{2}}.$$

I wish to thank you for your patience at this point and I promise you that it will pay off! In the next chapter we shall count the number of lattice points inside convex domains. This is one of the most interesting and difficult problems in modern mathematics. But before we do that... just a couple of exercises...

Exercise 11.7. Generalize Theorem 11.11 to any convex set in the plane whose boundary is infinitely differentiable and has non-vanishing curvature.

Exercise 11.8. (difficult) Generalize Theorem 11.11 to higher dimensions. What are you planning to use in place of Green's theorem? Just as importantly, what are you planning to use in place of van der Corput lemma?

I do not want to give you a hint here. By beating your head against the wall on this one, you will learn much...

1. Notes, remarks and difficult questions

Results of this chapter represent some special cases of the extensive and useful theory of the method of stationary phase. Most books where this method is presented require quite a bit of background to read, but there are some exceptions. Take a look at the presentation of these ideas in [**17**]. Even if you do not know what some terms mean, you will get a lot out of it.

One of the key tools in the higher dimensional version of the method of stationary phase is Morse's lemma. It says, roughly, that if the boundary of a convex body is smooth and has non-vanishing curvature, then after a change of variables, this boundary can be locally written as a graph of $x_1^2 + x_2^2 + \cdots + x_{d-1}^2$. What does non-vanishing curvature mean in higher dimensions? Here is a fancy but useful definition. Let N denote the Gauss map which maps each point of the boundary of a convex set to the unit sphere by taking each point to the unit normal at that point. The Gaussian curvature is defined as the determinant of the differential of this map. For a detailed description of these concepts, take a look at a beautiful textbook by Manfredo do Carmo ([**6**]).

Chapter 12

Integer points and a crash course on Fourier analysis

Finally! We get to apply the nasty estimates of the last chapter to something interesting and even exciting... A word of warning though. Something peculiar is being attempted in this chapter. I attempt to present the proof in the way that I thought about it the first time I looked at the problem. It is far from elegant, but my goal is to get across the messy nature of the process of discovery.

Let

$$N(t) = \#\{tD \cap \mathbb{Z}^2\},$$

where D, as before, is the unit disk in the plane. In other words, $N(t)$ counts the numbers of points with integer coordinates that are contained in the disk of radius t. How many such points are there? Well, it is intuitively clear that

$$N(t) \approx \pi t^2,$$

whatever $\approx$ means... After all, integration in two dimensions is sort of based on this principle, is it not? Based on this idea, we write

$$N(t) = \pi t^2 + E(t).$$

Now, we can write identities like this all day and night, but they do not mean anything unless $E(t)$ is actually much smaller than the "main" term πt^2. This is what this chapter is about and while the details will get very nasty, keep your eye on the eventual resolution of the lattice point question.

Exercise 12.1. Prove that there exists $C > 0$ such that

$$|E(t)| \leq Ct. \tag{12.1}$$

Hint: Draw a cube of side-length one centered at every point of $\mathbb{Z}^2$ and go to work...

Our goal is to beat (12.1). It was conjectured by Hardy, a great British mathematician of the first part of the 20th century, that for every $\epsilon > 0$ there exists C_ϵ such that

$$|E(t)| \leq C_\epsilon t^{\frac{1}{2}+\epsilon}.$$

He also proved that such a result would be the best possible, but that argument is outside the scope of this book. Hardy's conjecture is far out of reach at the moment, with the best known result being close to $Ct^{\frac{5}{8}}$, but this still constitutes tremendous progress from (12.1)!

We shall prove the following theorem established by W. Sierpinski in 1903.

Theorem 12.2. *With the notation above,*

$$|E(t)| \leq Ct^{\frac{2}{3}}.$$

To pay homage to Hardy's conjecture, we prove the following "average" version.

Theorem 12.3. *Let*

$$N_x(t) = \#\{(tD - x) \cap \mathbb{Z}^d\}.$$

Then

$$N_x(t) = \pi t^2 + E_x(t),$$

where

$$\left(\int_0^1 \int_0^1 |E_x(t)|^2 dx_1 dx_2 \right)^{\frac{1}{2}} \leq Ct^{\frac{1}{2}}. \tag{12.4}$$

Our proof of Theorem 12.2 will be closer to the one used by E. Landau, one of the greatest number theorists of all time, to prove a higher dimensional version of the same result. Several aspects of the proof are difficult to motivate, but we will try. We also hope that the sheer beauty of the argument will compensate for several rough moments that will come along the way.

What is $N(t)$? By definition, it is

$$\sum_{n\in\mathbb{Z}^2} \chi_{tD}(k).$$

This is a function of a single variable t, which is slightly unsettling, perhaps, because we have two-dimensional quantities floating about. Let's make $N(t)$ into a two-dimensional function

$$N_x(t) = \sum_{n\in\mathbb{Z}^2} \chi_{tD}(x-n).$$

What is this? Well, $N_x(t)$ measures the number of lattice points inside the disk of radius t centered at $x \in \mathbb{R}^2$. Moreover, this function is periodic in the sense that

$$N_{x+m}(t) = N_x(t) \tag{12.5}$$

for any $m \in \mathbb{Z}^2$.

Exercise 12.2. Carefully verify (12.5) by writing out the definition.

What other periodic functions do we know? There are trigonometric functions like sines and cosines, of course. Hmmm... what if we were to explore a naive notion that all periodic functions are somehow related. More precisely, given that DeMoivre's formula tells us that $e^{i\theta} = \cos(\theta) + i\sin(\theta)$, we could try to write a periodic function $F(x)$, like $N_x(t)$, for example, in the form

$$\sum_{k\in\mathbb{Z}^2} c_k e^{2\pi i x\cdot k} \tag{12.6}$$

for $x \in [0,1]^2$.

Suppose for a moment that we can in the sense that $F(x)$ equals the expression in (12.6). Observe that for $m \in \mathbb{Z}^2$, we then have

$$\int_{[0,1]^2} e^{-2\pi i x\cdot m} F(x)dx = \int_{[0,1]^2} e^{-2\pi i x\cdot m} \sum_{k\in\mathbb{Z}^2} e^{2\pi i x\cdot k} c_k$$

$$= \sum_{k\in\mathbb{Z}^2} c_k \int_{[0,1]^2} e^{-2\pi i x\cdot(m-k)} dx = c_m,$$

since

$$\int_{[0,1]^2} e^{-2\pi i x\cdot(m-k)} dx = 0 \tag{12.7}$$

if $k \neq m$ and 1 otherwise.

What did we just prove? Let us record it as a lemma.

Lemma 12.8. *Suppose that $F(x)$ is a periodic function and*

$$F(x) = \sum_{k\in\mathbb{Z}^2} c_k e^{2\pi i x\cdot k},$$

such that

$$\sum_k |c_k| < \infty.$$

Then

$$c_m = \int_{[0,1]^2} e^{-2\pi i x\cdot m} F(x)dx.$$

We shall henceforth refer to c_ms as Fourier coefficients of F... Yes yes... there is theory behind all this... but I am not going to tell you about it for this is not my goal!

Thus, if we could write

$$N_x(t) = \sum_{k\in\mathbb{Z}^2} c_k(t) e^{2\pi i x\cdot k},$$

then

$$c_m(t) = \int_{[0,1]^2} e^{-2\pi i x\cdot m} N_x(t)dx$$

$$= \int_{[0,1]^2} e^{-2\pi i x\cdot m} \sum_{n\in\mathbb{Z}^2} \chi_{tD}(x-n)dx$$

$$= \sum_{n\in\mathbb{Z}^2} \int_{[0,1]^2} e^{-2\pi i x\cdot m} \chi_{tD}(x-n)dx$$

$$
\begin{aligned}
&= \sum_{n\in\mathbb{Z}^2} \int_{[0,1]^2-n} e^{-2\pi i x\cdot m} e^{2\pi i n\cdot m} \chi_{tD}(x)dx \\
&= \int_{\mathbb{R}^2} e^{-2\pi i x\cdot m} \chi_{tD}(x)dx \\
&= \int_{\mathbb{R}^2} e^{-2\pi i x\cdot m} \chi_D(x/t)dx \\
&= t^2 \int_{\mathbb{R}^2} e^{-2\pi i x\cdot tm} \chi_D(x)dx \\
&= t^2 \widehat{\chi}_D(tm), \qquad (12.9)
\end{aligned}
$$

where the last step uses a definition from the end of the previous chapter.

We just discovered that if $N_x(t)$ can be written in the form

$$\sum_{k\in\mathbb{Z}^2} e^{2\pi i x\cdot k} c_k,$$

then c_k is given by the formula in Lemma 12.8 above. Now, let

$$N_x^*(t) = t^2 \sum_{m\in\mathbb{Z}^2} e^{2\pi i x\cdot m} \widehat{\chi}_D(tm).$$

If we can argue that the right hand side is actually a function, in some reasonable sense, then we will be close to showing that $N_x(t) = N_x^*(t)$ and we would have a beautiful formula

$$N_x(t) = \sum_{n\in\mathbb{Z}^2} \chi_{tD}(x-n) = t^2 \sum_{k\in\mathbb{Z}^2} e^{2\pi i x\cdot m} \widehat{\chi}_D(tm), \qquad (12.10)$$

and

$$N(t) = t^2 \sum_{k\in\mathbb{Z}^2} \widehat{\chi}_D(tm). \qquad (12.11)$$

There is much we can do with these formulas, but we need to clean up some details first. To begin with, observe that

$$
\begin{aligned}
\int_{[0,1]^2} |N_x^*(t) - \pi t^2|^2 dx &= \int_{[0,1]^2} \left| t^2 \sum_{m\in\mathbb{Z}^2} e^{2\pi i x\cdot m} \widehat{\chi}_D(tm) - \pi t^2 \right|^2 dx \\
&= t^4 \int_{[0,1]^2} \sum_{m\neq(0,0)} \sum_{m'\neq(0,0)} e^{2\pi i x\cdot(m-m')} \widehat{\chi}_D(tm)\overline{\widehat{\chi}_D(tm')} dx
\end{aligned}
$$

$$= t^4 \sum_{m\neq(0,0)} \sum_{m'\neq(0,0)} \widehat{\chi}_D(tm)\overline{\widehat{\chi}_D(tm')} \int_{[0,1]^2} e^{2\pi i x\cdot(m-m')} dx$$

$$= t^4 \sum_{m\neq 0} |\widehat{\chi}_D(tm)|^2$$

$$\leq Ct^4 t^{-3} \sum_{m\neq(0,0)} |m|^{-3} \leq C't, \tag{12.12}$$

since

$$\sum_{m\neq(0,0)} |m|^{-3}$$

converges by the integral test (check!) and

$$|\widehat{\chi}_D(tm)| \leq C(t|m|)^{-\frac{3}{2}}$$

by Theorem 11.11.

Exercise 12.3. There are missing steps above... For example, how did subtracting πt^2 result in summing over $m \neq (0,0)$. Just check everything...

The estimate (12.12) proves Theorem 12.3 provided that $N_x(t)$ actually equals $N_x^*(t)$, which we have not quite demonstrated. What do we have here? Well, there are two functions, $N_x(t)$ and $N_x^*(t)$. The mth Fourier coefficients of $N_x(t)$ is $t^2\widehat{\chi}_D(tm)$, by (12.9) above. We also know, by construction, that the mth Fourier coefficient of $N_x^*(t)$ is $t^2\widehat{\chi}_D(tm)$. It remains for us to argue that these two functions must in fact be the same. Well, we know that Fourier coefficients of $N_x(t) - N_x^*(t)$ are all zero. This implies immediately that

$$\int_{[0,1]^2} (N_x(t) - N_x^*(t))P(x)dx = 0$$

for every trigonometric polynomial $P(x)$, by which we mean a function of the form

$$P(x) = \sum_{k\in S\subset\mathbb{Z}^2} a_k e^{2\pi i x\cdot k},$$

where a_ks are complex numbers and S is finite. In Exercise 12.7 at the end of this chapter we shall outline a proof of the fact that this implies that in fact $N_x(t) = N_x^*(t)$. Taking this for granted for the moment, we conclude that in light of (12.12), Theorem 12.3 is proved.

We now turn our attention to Theorem 12.2 and we begin by deliberately failing! Setting $x = (0,0)$ we can argue that

$$N(t) = N_{(0,0)}(t) = N^*_{(0,0)}(t) = t^2 \sum_{m \in \mathbb{Z}^2} \widehat{\chi}_D(tm)$$

$$= \pi t^2 + t^2 \sum_{m \neq (0,0)} \widehat{\chi}_D(tm),$$

which leads us, among other things, to the conclusion that

$$E(t) = t^2 \sum_{m \neq (0,0)} \widehat{\chi}_D(tm).$$

We can then bravely press along and apply Theorem 11.11, which leads to

$$|E(t)| \leq t^2 \sum_{m \neq (0,0)} |\widehat{\chi}_D(tm)| \leq t^{\frac{1}{2}} \sum_{m \neq (0,0)} |m|^{-\frac{3}{2}}.$$

This is somewhat problematic since $\sum_{m \neq (0,0)} |m|^{-\frac{3}{2}}$ diverges... What do we do? Summing all the way to infinity does not appear to be profitable, but there is hope. Do you see it? On one hand we fail because we get a divergent series. On the other hand, the term in front of the divergent series is $t^{\frac{1}{2}}$, while we are perfectly happy to settle for $t^{\frac{2}{3}}$. With this in mind, define

$$N_R(t) = t^2 \sum_{m \neq (0,0)} \widehat{\chi}_D(tm) \pi^{-1} \widehat{\chi}_D(m/R).$$

What made us think of such a monster, you ask... Well, if there is one thing we learned from this book is that "you've got to pay the piper...". We could just cut off the summation at $|m| \leq R$, but then we still need to interpret what N_R actually means... This will presumably involve undoing the "$\widehat{\ }$"symbol... With that in mind, we may as well deal with the "$\widehat{\ }$" quantity now in hopes of looking at something less perverse in the process of understanding the nature of $N_R(t)$. We need $\widehat{\chi}_D(m/R)$ to play roughly the same roll as putting in the cut-off $|m| \leq R$. Let us see if it does.

We have

$$N_R(t) = \pi t^2 + E_R(t),$$

where

$$E_R(t) = t^2 \sum_{m \neq (0,0)} \widehat{\chi}_D(tm)\pi^{-1}\widehat{\chi}_D(m/R)$$

$$= t^2 \sum_{1 \leq |m| \leq R} \widehat{\chi}_D(tm)\pi^{-1}\widehat{\chi}_D(m/R)$$

$$+t^2 \sum_{|m|>R} \widehat{\chi}_D(tm)\pi^{-1}\widehat{\chi}_D(m/R) = I + II.$$

Applying Theorem 11.11 once again, and observing that

$$\pi^{-1}|\widehat{\chi}_D(m/R)| \leq 1 \tag{12.13}$$

we see that

$$I \leq Ct^{\frac{1}{2}} \sum_{1 \leq |m| \leq R} |m|^{-\frac{3}{2}} \leq C' t^{\frac{1}{2}} R^{\frac{1}{2}},$$

which is great, but we need to worry about II...

Exercise 12.4. Verify (12.13) by observing, in general, that

$$|\widehat{f}(\xi)| = \left| \int_{\mathbb{R}^d} e^{-2\pi i x \cdot \xi} f(x) dx \right|$$

$$\leq \int_{\mathbb{R}^d} |f(x)| dx.$$

We have, again by Theorem 11.11,

$$II \leq Ct^{\frac{1}{2}} \sum_{|m|>R} |m|^{-3} R^{\frac{3}{2}} \leq C' t^{\frac{1}{2}} R^{\frac{1}{2}}.$$

We can thus conclude that

$$|E_R(t)| \leq Ct^{\frac{1}{2}} R^{\frac{1}{2}}. \tag{12.14}$$

Our problem, of course, is that we cannot have R in our estimates. This simply means that we need another estimate on $N_R(t)$ that we can use to weed the influence of R out. If we could somehow conclude that we may take $R = t^{\frac{1}{3}}$, the proof of Theorem 12.2 would instantly follow!

What is N_R? We could go through the headache we encountered in the proof of Theorem 12.3, but we do not need to. Observe the following simple manipulation. Let f and g be functions on $\mathbb{R}^2$. Define

$$f * g(x) = \int_{\mathbb{R}^2} f(x-y)g(y)dy,$$

and consider

$$\widehat{f * g}(\xi) = \int e^{-2\pi i x\cdot\xi} \int_{\mathbb{R}^2} f(x-y)g(y)dydx$$

$$= \int\int e^{-2\pi i(x-y)\cdot\xi} e^{-2\pi i y\cdot\xi} f(x-y)g(y)dydx$$

$$= \int\int e^{-2\pi i u\cdot\xi} e^{-2\pi i v\cdot\xi} f(u)g(v)dudv = \widehat{f}(\xi)\widehat{g}(\xi).$$

With this observation in hand we think about what we learned from the proof of Theorem 12.3. In the game we played with Fourier coefficients it did not really matter that we were working with the specific function $N_x(t)$, right? What we actually demonstrated is that if ϕ is a sufficiently well-behaved function, then

$$\sum_{n\in\mathbb{Z}^2} \phi(n) = \sum_{k\in\mathbb{Z}^2} \widehat{\phi}(k), \tag{12.15}$$

where

$$\widehat{\phi}(\xi) = \int_{\mathbb{R}^2} e^{-2\pi i x\cdot\xi}\phi(x)dx,$$

the Fourier transform of ϕ. This is known as the Poisson Summation Formula (PSF) and constitutes a high point of the elementary Fourier transform theory.

What does this tell us about $N_R(t)$? Well, by (12.15),

$$N_R(t) = t^2 \sum_{m\in\mathbb{Z}^2} \widehat{\chi}_D(tm)\pi^{-1}\widehat{\chi}_D(m/R)$$

$$= R^2 \sum_{n\in\mathbb{Z}^2} \chi_{tD} * \chi_{R^{-1}D}(n). \tag{12.16}$$

Exercise 12.5. Verify by a direct calculation that the Fourier transform of $t^2\widehat{\chi}_D(t\cdot)$ is indeed $\chi_{tD}(\cdot)$ and the Fourier transform of $\widehat{\chi}_D(\cdot/R)$ is $R^2\chi_{R^{-1}D}(\cdot)$.

With (12.16) in tow, it is not hard to see that

$$N(t) \leq N_R(t + R^{-1}). \tag{12.17}$$

Why is this not hard, you ask? Think about the function $V(x) = R^2 \chi_{tD} * \chi_{R^{-1}D}(x)$. If $|x| > t + R^{-1}$, $V(x) \equiv 0$. If $x \leq t + R^{-1}$, then $V(x) \geq 1$. The inequality (12.17) follows immediately.

Exercise 12.6. Check the assertions of the previous paragraph by a direct calculation.

We are now ready for the punch line! We have

$$N(t) = \pi t^2 + E(t),$$

and

$$N_R(t + R^{-1}) = \pi(t + R^{-1})^2 + E_R(t + R^{-1}).$$

Using these identities, (12.16) and (12.14), we see that

$$|E(t)| \leq C(tR^{-1} + |E_R(t)|). \tag{12.18}$$

Combining (12.14) with (12.18) and taking $R = t^{\frac{1}{3}}$, we see that

$$|E(t)| \leq Ct^{\frac{2}{3}},$$

and Theorem 12.2 is proved.

Exercise 12.7. Prove that if $\int_{[0,1]^2} |F(x)|^2 dx < \infty$ and all the Fourier coefficients of F are 0, then $F(x) \equiv 0$ by going through the following outline.

Let

$$P_n(x) = \frac{1}{(n+1)^2} \frac{\sin^2(\pi(n+1)x_1)}{\sin^2(\pi x_1)} \frac{\sin^2(\pi(n+1)x_2)}{\sin^2(\pi x_2)}.$$

Prove that

$$\int_{[0,1]^2} P_n(x)dx = 1$$

and use it to prove that

$$\int_{[0,1]^2} |F * P_n(x) - F(x)|^2 dx \to 0 \text{ as } n \to \infty. \tag{12.19}$$

Prove that $F * P_n(x)$ is a trigonometric polynomial by showing that $P_n(x)$ is a trigonometric polynomial and that

$$F * P_n(x) = \sum_{k \in \mathbb{Z}^2} c_k \widehat{P}_n(k),$$

where c_ks are the Fourier coefficients of F.

Conclude using (12.19) that

$$\int_{[0,1]^2} |F(x)|^2 dx = 0,$$

which implies that $F(x) \equiv 0$ since $|F(x)|^2 \geq 0$.

Exercise 12.8. Prove the higher dimensional analog of the main result of this chapter. More precisely, show that if

$$N(t) = \#\{tB_d \cap \mathbb{Z}^d\},$$

where B_d is the unit ball in $\mathbb{R}^d$. Prove that

$$N(t) = |B_d| t^d + E(t),$$

where

$$|E(t)| \leq C t^{d-2+\frac{2}{d+1}}.$$

Hint: You will need one of the exercises from the previous chapter.

Exercise 12.9. (difficult) Replace the disk in this chapter by the set

$$D_m = \{x \in \mathbb{R}^2 : x_1^m + x_2^m \leq 1\},$$

where m is, say, an even integer greater than two. Prove an estimate of the form

$$N(t) = t^2 |D_m| + E(t),$$

with

$$|E(t)| \leq C t^{\gamma}.$$

Obtain the smallest γ you can... Can you get $\gamma = \frac{m-1}{m}$? Burton Randol did... Again, I could give you a hint, but I do not want to. You will first come up with a non-trivial but bad exponent, then you will think harder...

Exercise 12.10. With the notation of this chapter, prove that

$$\left(R^{-1} \int_R^{2R} |E(t)|^2 dt \right)^{\frac{1}{2}} \leq C\sqrt{R}.$$

Hint: Use the fact, not proved in this book, that

$$\widehat{\chi}_D(\xi) = C|\xi|^{-\frac{3}{2}} \cos\left(2\pi\left(|\xi) - \frac{1}{8}\right)\right) + O(|\xi|^{-\frac{5}{2}}.$$

Take a look at [**11**] for a more general version of this problem. By the way, can you go ahead and prove that for every $\epsilon > 0$ there exists $C_\epsilon > 0$ such that

$$\left(R^{-1} \int_R^{2R} |E(t)|^4 dt \right)^{\frac{1}{4}} \leq C_\epsilon R^{\frac{1}{2}+\epsilon}.$$

In the process of trying to solve this problem you will encounter the question of how many lattice points can live in certain circular annuli. I will not tell you anything about that, but you can look it up on the web..

Exercise 12.11. Speaking of annuli, use the results of this chapter to show that

$$\#\{m \in \mathbb{Z}^2 : R \leq |m| \leq R + h\} \geq CRh,$$

provided that

$$h \geq cR^{-\frac{1}{3}}.$$

Hint: Observe that

$$\#\{m \in \mathbb{Z}^2 : R \leq |m| \leq R + h\} = N(R + h) - N(R).$$

1. Notes, remarks and difficult questions

As we mention above, the Holy Grail of this subject is to prove the Hardy Conjecture:

$$|E(t)| \leq Ct^{\frac{1}{2}+\epsilon}$$

for any $\epsilon > 0$.

The best current result has $\frac{1}{2}$ replaced by a number just bigger than $\frac{5}{8}$ and no end is in sight. In fact, the author of this book, in a fit of bad humor, once plotted the results related to this problem

from 1903 until the present and projected that the Hardy Conjecture will be proved sometime in the 24th century... There is no need to comment on the absurdity of this "analysis", but it does indicate how hard people have worked on this problem without proving the ultimate conjecture.

In three dimensions, the situation is similarly complicated. It is conjectured that

$$|E(t)| \leq Ct^{1+\epsilon}.$$

The best known result, due to Heath-Brown, is

$$|E(t)| \leq Ct^{\frac{19}{15}},$$

and the ultimate solution seems years away.

Somewhat surprisingly, the problem is solved in higher dimensions. It is know that if $d \geq 5$,

$$|E(t)| \leq Ct^{d-2},$$

and if $d = 4$,

$$|E(t)| \leq Ct^2 \log^2(t).$$

In both cases the result cannot be improved, except possibly by a logarithm in four dimensions. The fact that the higher dimensional case turns out to be easier, is related, though not limited to, the beautiful fact that if $d \geq 5$, and R is a square root of an integer, then there exist $C_1, C_2 > 0$ such that

$$C_1 R^{d-2} \leq \#\{RS^{d-1} \cap \mathbb{Z}^d\} \leq C_2 R^{d-2},$$

a remarkably stable estimate.

Chapter 13

Return of the Fourier transform

After surviving the wrestling match that is the previous chapter, we are surely ready for something less technical yet equally interesting. Remember $\mathbb{F}_q$, the finite field with q elements? You do not? Well, then go and read Chapter 5 and come back later :). Having motivated the Fourier transform in the previous chapter, I feel fully justified in defining it on $\mathbb{F}_q$, and indeed on $\mathbb{F}_q^d$, without making any further excuses.

Let $f : \mathbb{F}_q^d \to \mathbb{C}$, where q is a prime. Define for every $m \in \mathbb{F}_q^d$,

$$\widehat{f}(m) = q^{-d} \sum_{x \in \mathbb{F}_q^d} \chi(-x \cdot m) f(x),$$

where

$$\chi(t) = e^{-\frac{2\pi i}{q} t},$$

and, as usual, $x = (x_1, \ldots, x_d)$.

Want to know why the Fourier transform on $\mathbb{F}_q^d$ is beautiful? Watch this... Consider

$$\sum_{m \in \mathbb{F}_q^d} \chi(x \cdot m) \widehat{f}(m)$$

$$= \sum_{m \in \mathbb{F}_q^d} \chi(x \cdot m) q^{-d} \sum_{y \in \mathbb{F}_q^d} \chi(-y \cdot m) f(y)$$

$$= q^{-d} \sum_{y \in \mathbb{F}_q^d} f(y) \sum_{m \in \mathbb{F}_q^d} \chi((x-y) \cdot m). \tag{13.1}$$

Exercise 13.1. By computing the geometric series, show that

$$\sum_{m \in \mathbb{F}_q^d} \chi((x-y) \cdot m) = q^d \delta(x-y),$$

where $\delta(s) = 1$ if $s = 0$ and 0 otherwise.

Using Exercise 13.1 we see that (13.1) takes the form

$$\sum_{y \in \mathbb{F}_q^d} f(y) \delta(x-y) = f(x)$$

by definition of δ.

We have just proved the Fourier Inversion Formula on $\mathbb{F}_q^d$:

Lemma 13.2. *We have*

$$\sum_{m \in \mathbb{F}_q^d} \chi(x \cdot m) \widehat{f}(m) = f(x).$$

13.1. Let's play a little more. What other treasures does this discrete Fourier transform have in store for us? As we have nothing better to do, let us consider a sum

$$\sum_{m \in \mathbb{F}_q^d} \widehat{f}(m).$$

By the Fourier inversion formula, this expression equals $f(0, \ldots, 0)$, which is interesting. Let us keep going. What about

$$\sum_{m \in \mathbb{F}_q^d} |\widehat{f}(m)|?$$

Unfortunately, the presence of the absolute value prevents us from doing anything nearly as nice as before. We could just estimate this quantity head on, however. We have

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

$$\leq q^{-d} \sum_{x \in \mathbb{F}_q^d} |f(x)|,$$

since $|\chi(t)| \equiv 1$. It follows that

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

In particular, let $E \subset \mathbb{F}_q^d$ and define $E(x)$ to equal to 1 if $x \in E$ and 0 otherwise. We shall henceforth refer to $E(x)$ as the indicator function of E. Then, by above,

$$\sum_{m \in \mathbb{F}_q^d} |\widehat{E}(m)| \leq \sum_{x \in \mathbb{F}_q^d} E(x) = \#E.$$

Let's keep moving. What about

$$\sum_{m \in \mathbb{F}_q^d} |\widehat{f}(m)|^2?$$

One thing we are good at after reading this book is squaring things. We have

$$\sum_{m \in \mathbb{F}_q^d} |\widehat{f}(m)|^2 = \sum_{m \in \mathbb{F}_q^d} q^{-2d} \sum_{x,y \in \mathbb{F}_q^d} \chi(-(x-y) \cdot m) f(x) \overline{f(y)}$$

$$q^{-2d} \sum_{x,y \in \mathbb{F}_q^d} \sum_{m \in \mathbb{F}_q^d} \chi(-(x-y) \cdot m) f(x) \overline{f(y)}$$

$$= q^{-2d} \sum_{x,y \in \mathbb{F}_q^d} q^d \delta(x-y) f(x) \overline{f(y)}$$

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

and we have proved the Plancherel theorem:

Theorem 13.3. *Let $f : \mathbb{F}_q^d \to \mathbb{C}$. Then*

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

13.2. What information does the Fourier transform encode? Take $E \subset \mathbb{F}_q^d$ and take a look at

$$\sum_{m\in\mathbb{F}_q^d} |\widehat{E}(m)|^4 = q^{-4d} \sum_{x,y,x',y'\in\mathbb{F}_q^d} E(x)E(y)E(x')E(y')$$

$$\times \sum_{m\in\mathbb{F}_q^d} \chi((x+y-x'-y')\cdot m)$$

$$= q^{-3d} \sum_{x,y,x',y'\in\mathbb{F}_q^d} E(x)E(y)E(x')E(y')\delta(x+y-x'-y')$$

$$= q^{-3d}\#\{(x,y,x',y')\in E^4 : x+y=x'+y'\}.$$

Exercise 13.2. Prove that

$$\#\{(x,y,x',y')\in E^4 : x+y=x'+y'\} \leq C(\#E)^3.$$

Can you compute the best possible constant C?

In other words,

$$\#\{(x,y,x',y')\in E^4 : x+y=x'+y'\} = q^{3d} \sum_{m\in\mathbb{F}_q^d} |\widehat{E}(m)|^4. \tag{13.4}$$

We thus see that if $\widehat{E}(m)$ is small, then

$$\#\{(x,y,x',y')\in E^4 : x+y=x'+y'\}$$

is also small.

How small can $\widehat{E}(m)$ be? Well, we know from Theorem 13.3 that

$$q^{-d} \sum_{m\in\mathbb{F}_q^d} |\widehat{E}(m)|^2 \tag{13.5}$$

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

It follows that the best estimate we can ever hope for is

$$|\widehat{E}(m)| \leq Cq^{-d}\sqrt{\#E}, \tag{13.6}$$

for suppose that

$$|\widehat{E}(m)| \leq \epsilon q^{-d}\sqrt{\#E}.$$

Then plugging this back into (13.5) we get

$$q^{-2d}\#E \leq q^{-d} \sum_{m \in \mathbb{F}_q^d} \left|\widehat{E}(m)\right|^2$$

$$\leq q^{-2d}\epsilon \#E.$$

It follows that if $\epsilon < 1$, we have a contradiction.

Now that we have an idea about how small $\widehat{E}(m)$ can be, it is reasonable to see what happens in (13.4) if the Fourier transform is this small. Suppose that (13.6) holds. Then

$$\#\{(x,y,x',y') \in E^4 : x+y = x'+y'\}$$

$$\leq Cq^{3d} \cdot q^d q^{-4d} (\#E)^2 = C(\#E)^2.$$

This is as good an estimate as we can hope for because

$$\#\{(x,y,x',y') \in E^4 : x+y = x'+y'\} \geq (\#E)^2$$

due to the fact that

$$\{(x,y,x,y) \in E^4\} \subset \{(x,y,x',y') \in E^4 : x+y = x'+y'\}.$$

A reasonable thing to do at this point is to plug in the estimate (13.6) into (13.4) and see what happens. We get

$$\#\{(x,y,x',y') \in E^4 : x+y = x'+y'\}$$

$$\leq Cq^{3d} \cdot q^d \cdot \left(q^{-d} \cdot \sqrt{\#E}\right)^4 \leq C(\#E)^2.$$

What does this show? It shows that the best possible upper bound, up to a constant, on $\#\{(x,y,x',y') \in E^4 : x+y = x'+y'\}$ is achieved if we have the best possible upper bound on $|\widehat{E}(m)|$. We still need to make sure that we are talking about something that actually exists! More precisely, how do we know that sets that satisfy (13.6) actually exist?! This is the subject of the next subsection.

13.3. Small Fourier transform. Let us get down to actually exhibiting a family of sets whose Fourier transform is as small as the law allows. Let

$$E = \{x \in \mathbb{F}_q^d : x_d = x_1^2 + \cdots + x_{d-1}^2\}.$$

Then

$$\widehat{E}(m) = q^{-d} \sum_{y \in \mathbb{F}_q^{d-1}} \chi(-y_1 m_1 - \cdots - y_{d-1} m_{d-1} - m_d(y_1^2 + \cdots + y_{d-1}^2)).$$

By separating variables, it suffices to consider

$$F(m_j, m_d) = \sum_{t \in \mathbb{F}_q} \chi(-tm_j - t^2 m_d)$$

$$= \sum_{t \in \mathbb{F}_q} \chi\left(-m_d\left(t + \frac{m_j}{2m_d}\right)^2 + \frac{m_j^2}{4m_d}\right)$$

$$= \chi\left(\frac{m_j^2}{4m_d}\right) \sum_{t \in \mathbb{F}_q} \chi(-m_d t^2).$$

Observe that if $m_d = 0$, $F(m_1, 0) \equiv 0$, unless $m_1 = 0$. We are now faced with the problem of estimating sums of the form $\sum_{t \in \mathbb{F}_q} \chi(-m_d t^2)$. How do we do this? If we square the sum, we get

$$\left| \sum_{t \in \mathbb{F}_q} \chi(-m_d t^2) \right|^2 = \sum_{t,s \in \mathbb{F}_q} \chi(-m_d(t^2 - s^2))$$

$$= \sum_{u \in \mathbb{F}_q} \chi(-m_d u) n(u),$$

where

$$n(u) = \#\{(t,s) \in \mathbb{F}_q \times \mathbb{F}_q : t^2 - s^2 = u\}$$

$$= \#\{(t,s) \in \mathbb{F}_q \times \mathbb{F}_q : ts = u\}.$$

Exercise 13.3. Verify all the details of the alleged equality in the last two lines.

If $u = 0$, $n(u) = 2q - 1$ since one of t and s must be zero and the other can take on whatever value. If $u \neq 0$, $n(u) = q - 1$ since once t is chosen, say, s is completely determined by the equation $s = ut^{-1}$, where inversion is taken modulo q. It follows that

$$\sum_{u \in \mathbb{F}_q} \chi(-m_d u) n(u) = 2q - 1 + (q-1) \sum_{u \neq 0} \chi(-m_d u)$$

$$= q + (q-1) \sum_{u \in \mathbb{F}_q} \chi(-m_d u) = q.$$

Exercise 13.4. Check this last step...

We conclude that

$$|F(m_1, m_d)|^2 = q,$$

if $m_d \neq 0$ and thus

$$|\widehat{E}(m)| \leq q^{-d} \cdot q^{-\frac{d-1}{2}} = q^{-\frac{d+1}{2}}$$

provided that $m \neq (0, \ldots, 0)$.

Observe that since $\#E = q^{d-1}$ in this case, we may conclude that (13.6) is satisfied.

Combining everything we have done, we have the following theorem.

Theorem 13.7. *Let $E = \{x \in \mathbb{F}_q^d : x_d = x_1^2 + \cdots + x_d^2\}$. Then*

$$\#\{(x, y, x', y') \in E^4 : x + y = x' + y'\} \leq (\#E)^2. \tag{13.8}$$

Exercise 13.5. We have proved the estimate $|\widehat{E}(m)| \leq Cq^{-d}\sqrt{\#E}$, with E as in Theorem 13.7, if $m \neq (0, \ldots, 0)$. When $m = (0, \ldots, 0)$, $\widehat{E}(m) = q^{-d} \sum_{x \in \mathbb{F}_q^d} E(x) = q^{-d} \#E$. Go through the argument above carefully and make sure that the value at the origin does not ruin the proof.

Exercise 13.6. We made an assertion above that if E is as in Theorem 13.7, then $\#E = q^{d-1}$. Verify this assertion rigorously.

Exercise 13.7. Let $E = \{(t, t, \ldots, t) \in \mathbb{F}_q^d : t \in \mathbb{F}_q\}$. Compute $\widehat{E}(m)$ and show that the estimate (13.6) is not in general satisfied.

Exercise 13.8. (more of an exploration). Let $SO(2)$ denote the set of two by two matrices with determinant one and such that if the first row is (v_1, v_2) and the second row is (w_1, w_2), then $v_1 w_1 + v_2 w_2 = 0$. For each $T \in SO_2(\mathbb{F}_q)$, define $E_T(x) = E(Tx)$ and let $m \neq (0, \ldots, 0)$. What can you tell me about

$$\left(\frac{1}{\#SO_2(\mathbb{F}_q)} \sum_{T \in SO_2(\mathbb{F}_q)} |\widehat{E}_T(m)|^2 \right)^{\frac{1}{2}} ?$$

Exercise 13.9. Can you show that the conclusion of Theorem 13.7 holds for a "typical" set E? More precisely, let $\mathbb{S}_k$ denote the set of subsets of $\mathbb{F}_q^d$ of size k. Can you show that for a positive proportion of sets in $\mathbb{S}_k$, (13.8) holds.

1. Notes, remarks and difficult questions

Perhaps the most interesting set satisfying the estimate (13.6) is the discrete sphere

$$S_j = \{x \in \mathbb{F}_q^d : x_1^2 + x_2^2 + \cdots + x_d^2 = j\},$$

with $j \neq 0$.

Assume that

$$\left| \sum_{t \in \mathbb{F}_q^*} \chi(at + t^{-1}) \psi(t) \right| \leq 2\sqrt{q} \tag{13.9}$$

for any $a \neq 0$ and any multiplicative function ψ from $\mathbb{F}_q^*$ to the unit circle in the complex plane. Can you prove using this fact that $\widehat{S}_j(m)$, $m \neq (0, \ldots, 0)$, satisfies the estimate (13.6)? You may want to first check that

$$\widehat{S}_j(m) = q^{-d} \sum_{x \in \mathbb{F}_q^d} q^{-1} \sum_{t \in \mathbb{F}_q} \chi(t(x_1^2 + \cdots + x_d^2 - j)) \chi(-x \cdot m),$$

and go from there.

The sum in (13.9) is known as the Kloosterman sum, discovered by Kloosterman early in the 20th century. The estimate (13.9) is due to Andrew Weil ([**20**]) as a consequence of his proof of the Riemann hypothesis for finite fields. Unfortunately, the paper is not incredibly

accessible without quite a bit of background in number theory and algebraic geometry, but it is something to keep in mind as you learn more and more mathematics in the future.

Another entertaining set satisfying the estimate (13.6) is the hyperbola

$$H_c = \{x \in \mathbb{F}_q^2 : x_1 x_2 = c\},$$

with $c \neq 0$.

Can you use (13.9) to prove that $\widehat{H}_c(m)$ satisfies the estimate (13.6)? If you can, go on and prove that if $E, F \subset \mathbb{F}_q^2$, then

$$\begin{aligned}&\#\{(x,y) \in E \times F : (x_1 - y_1)(x_2 - y_2) = c\}\\ &\leq C(q^{-1} \cdot \#E \cdot \#F + \sqrt{q \cdot \#E \cdot \#F}).\end{aligned}$$

It only gets harder from here. Can you keep going and prove that if $A \subset \mathbb{F}_q$,

$$A + A = \{a + a' : a, a' \in A\};\ A \cdot A = \{a \cdot a' : a, a' \in A\},$$

then

$$\max\{\#A + A, \#A \cdot A\} \geq \frac{(\#A)^{\frac{3}{2}}}{q^{\frac{1}{4}}},$$

provided that $q^{\frac{1}{2}} \leq \#A \leq Cq^{\frac{7}{10}}$?

If you get stuck, take a look at ([**9**]) where this result was recently proved.

Chapter 14

It is time to say goodbye

You must have noticed that on many occasions, we had to go through a bunch of nasty and tiresome calculations to reach conclusions which, I hope, you found beautiful and satisfying. In this regard, I would mention the following piece of wisdom by Maimonides, perhaps the greatest scholar of the Hebrew Bible: "Be convinced that, if man were able to reach the end without preparatory studies, such studies would not be preparatory but tiresome and utterly superfluous". Many undergraduate students I have met over the years separated things they learned into "techniques" and "ideas". The most successful ones eventually learned that one cannot exist without the other and that the border areas are both fluid and grey.

Let us summarize the story we just told. We first proved an innocent looking Cauchy-Schwartz inequality and used it to estimate things as seemingly diverse as the number of incidences of points and lines and the size of projections of sets in various dimensions. We continued our exploration of higher dimensional geometry by investigating the possibility of fitting the unit cube inside projections of balls in arbitrarily high dimension. From the study of properties of concrete sets we moved on to the Kakeya problem where we asked how small a set can be if it still contains a line in every possible direction in a vector space over a finite field. Note that this problem is about pathological sets, not typical ones, so in order not to become

too fascinated by weirdness, we moved on to probabilistic notions, both in discrete and continuous settings, culminating in a classical argument that the probability that two positive integers are relatively prime is $\frac{6}{\pi^2}$. Having built up some technique and tolerance for complicated arguments, we plunged into trigonometric integrals and their application to a beautiful problem of counting the number of lattice points inside large disks while in the process exploring properties of the Fourier transform in Euclidean space. Building upon the intuition we built up, we returned to finite fields in the final chapter and showed that size properties of the Fourier transform are intimately related to arithmetic properties of sets.

What should one do after having muscled his or her way through this labyrinth of calculations and ideas? One may be tempted to keep on reading and learning and that is a great thing. However, it is almost more important to use these ideas as a springboard for exploration and creative searching. The concepts presented in each chapter and even the exercises and comments at the end only give you a taste of the subject matter. Can you anticipate further developments? Can you formulate key questions that could lead to further progress? Are you willing to tirelessly search the research journals and the internet to find out what the concepts you have been introduced are connected to? This book can only be called a success if it causes you to do all these things and more. Good luck!

Bibliography

[1] T. M. Apostol, *Introduction to Analytic Number Theory*, New York: Springer-Verlag(1976).

[2] J. Pach and P. Agarwal, *Combinatorial geometry,* Wiley-Interscience Series in Discrete Mathematics and Optimization, A Wiley-Interscience Publication, John Wiley and Sons, Inc., New York (1995).

[3] J. Bourgain, N. Katz and T. Tao, *A sum-product estimate in finite fields, and applications*, Geom. Func. Anal **14** (2004) 27-57.

[4] A. Cordoba, *The Kakeya maximal function and spherical summation multipliers*, Amer. J. Math. **99** (1977), 1-22.

[5] R. Davies, *Some remarks on the Kakeya problem*, Proc. Camb. Phil. Soc. **69** (1971), 417-421.

[6] M. do Carmo, *Differential geometry of curves and surfaces*, Prentice Hall (1976).

[7] W. Feller, *Stirling's Formula* 2.9 in An Introduction to Probability Theory and Its Applications **1**, 3rd ed. New York: Wiley, 50-53, (1968).

[8] M. Huxley, *Area, Lattice Points, and Exponential Sums*, London Mathematical Society Monographs, **13**, Oxford University Press (1996).

[9] D. Hart, A. Iosevich and J. Solymosi, *Sum-product estimates in finite fields via Kloosterman sums*, (preprint), (2006).

[10] A. Iosevich, *Geometric measure theory and Fourier analysis*, Birkhauser; proceedings of the series of lectures delivered at Padova (Minicorsi) in 2002 (2004).

[11] A. Iosevich, E. Sawyer and A. Seeger *Mean square discrepancy bounds for the number of lattice points inside large convex domains*, Journal D'Analyse, **87**, (2002), 209-230.

[12] N. Katz and T. Tao, *Some connections between the Falconer and Furstenburg conjectures* , New York J. Math. **7** (2001), 148-187.

[13] N. Katz and G. Tardos, *A new entropy inequality for the Erdos distance problem*, Towards a Theory of Geometric Graphs. (ed.J Pach) Contemporary Mathematics **342** (2004).

[14] N. Katz, I. Laba, and T. Tao, *An improved bound on the Minkowski dimension of Besicovitch sets in* $\mathbb{R}^3$, Annals of Math. **152** (2000), 383-446.

[15] L.H. Loomis and H. Whitney, *An inequality related to the isoperimetric inequality*, Bull. Amer. Math. Soc. **55** (1949), 961-962.

[16] J. Steele, *The Cauchy-Schwartz Master Class*, Cambridge University Press, Cambridge, (2004).

[17] C. Sogge, *Fourier integrals in classical analysis*, Cambridge University Press, Cambridge, (1993).

[18] E. Szemeredi and W. Trotter, *Extremal problems in discrete geometry*, Combinatorica **3** (1983), 381-392.

[19] J. Solymosi and C. Toth, *Distinct distances in the plane*, Discr. Comp. Jour. (Misha Sharir birthday issue) **25** (2001), 629-634.

[20] A. Weil, *On some exponential sums*, Proc. Nat. Acad. Sci. U.S.A. **34** (1948), 204-207.

[21] T. Wolff, *Recent work connected with the Kakeya problem*, Prospects in Mathematics (Princeton, NJ, 1996) Amer. Math. Soc., Providence, RI (1999), 129-162.