Foreword

There are several goals for this book. As the title indicates, we certainly hope to familiarize you with some of the major results in the study of the Erdős distance problem. This goal should be easily attainable for most experienced mathematicians. However, if you are not an experienced mathematician, we hope to guide you through many advanced mathematical concepts along the way.

The book is based on the notes that were written for the summer program on the problem, held at the University of Missouri, August 1–5, 2005. This was the second year of the program, and our plan continued to be an introduction for motivated high school students to accessible concepts of higher mathematics.

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

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

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

Introduction

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

1. A sketch of our problem

In the case that there is only one point, we have but one distance, zero. It might seem odd to count zero as a distance, but it will make things easier later on if we just assume that it is. In the case of two points, our job is pretty easy again. We have the distance between the two points, and again, zero. However, if we consider the case of three points in the plane, it begins to get interesting. Three points arranged as the vertices of an equilateral triangle are the same distance from one another, so there is only one nonzero distance, making two total. If they are the vertices of an isosceles triangle, we have one distance repeated, leaving three distinct distances total. Of course, there are any number of ways for three points to determine four distances. These phenomena increase in complexity and frequency as we consider more and more points. In fact, there is no configuration of four points

in the plane that has only one nonzero distance present. It stands to reason that as we add more points, we will add more distances. To explore this problem, we fix a dimension to work in, d, and then investigate the *asymptotic* behavior which depends on the number of points, n, or how things happen as n grows large, past a million, past a billion, and so on. Since we are considering large n, we will not be concerned with the exact number of distinct distances, but with how many distinct distances there are in comparison to n.

In full generality, the Erdős distance problem asks for the minimum number of distances determined by n points in d-dimensional space, $\mathbb{R}^d$, where the minimum is taken over all the sets P containing n elements. For this to be interesting, we will assume that $d \geq 2$. In the case $d = 1$, it is easy to see that the number of distances determined by the set of n points is at least n, and this bound is achieved, for example, by the set $\{0, 1, \ldots, n-1\}$. When x is a point in d-dimensional space, we write its coordinates as $(x_1, x_2, \ldots, x_d)$. Define[2]

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

where

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

the standard Euclidean distance.

Using this notation, we want to know the smallest possible size of $\Delta(P)$ over all the sets P of a given fixed size. Let us consider some simple examples that involve many points. Let

$$P = \{(0,0), (1,0), \ldots, (n-1, 0)\}.$$

Then $\Delta(P) = \{0, 1, 2, \ldots, n-1\}$. This simple example shows that there is a set of n points that only determines exactly n distinct distances.

In general, we can construct a set of n points in $\mathbb{R}^d$, $d \geq 2$, that determine approximately $n^{\frac{2}{d}}$ distances when $d \geq 3$ and approximately

[2]Here, the colon next to the equals sign indicates that we are defining something. The colon inside the braces can be read as "such that". Here we are defining $\Delta(P)$ to be the set of distances, $|p - p'|$ **such that** p and p' are elements of P.

$\frac{n}{\sqrt{\log(n)}}$ distances when $d = 2$. This is achieved by taking all the points in the cube of side-length $n^{\frac{1}{d}}$, with sides parallel to the axes, having integer coordinates. It is not difficult to see that the number of distances determined by this set is $\lesssim n^{\frac{2}{d}}$ in every dimension. See Exercise 0.3 below. A bit of number theory is required for the lower bound and to establish the logarithmic loss in two dimensions. See [**12**] and the references contained therein.

These kinds of explorations are nowhere near to being fully understood, but much is known, and we will come very close to the cutting edge of this beautiful area of study in this book. One of the great things about this theory is that it can be developed largely from the ground up. That is, this problem in particular can be studied without much of background. So if you are curious as to what mathematical research is like, reading through this book can provide you with a glimpse. You can actively watch the theory grow from its infancy through some of the most recent discoveries in the field. Along the way, you will be introduced to many of the elementary techniques in any serious mathematician's toolkit. If you are already familiar with research mathematics, and desire more justification for serious exploration of this particular area, we have included sketches of some consequences of the study of this problem in the final chapter of this book. More precisely, we show how the Erdős distance problem and the Erdős integer distance principle can be used to demonstrate that a set of mutually orthogonal exponentials on a smooth symmetric convex surface in $\mathbb{R}^d$ with everywhere non-vanishing curvature must be very small. This provides a connection between a set of problems in classical analysis and the main theme of this book. This is just one of many connections between the Erdős distance problems and other areas of mathematics. An interested reader is encouraged to consult a beautiful article by Nets Katz and Terry Tao ([**27**]) and the references contained therein. See also [**16**].

2. Some notation

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

As above, if x is a vector, $|x| = \sqrt{x_1^2 + \cdots + x_d^2}$ will denote its (Euclidean) length, or distance from the origin. Of course, if y is also a vector, $|x - y|$ will denote the distance between x and y.

If A is a set, we can indicate the elements in the set as $A := \{a_1, a_2, \ldots, a_n\}$. We can designate the size of the set as $|A|$, or sometimes as $\#A$. Union and intersection are denoted as usual, with $\cup$ and $\cap$, respectively. If B is another set, we use $A \setminus B$ to mean all of the elements in A that are not in B. We write the *Cartesian product* of A and B as $A \times B$. It is defined as the set of all pairs of elements, (a, b), where $a \in A$, and $b \in B$.

Consider two sets, $A := \{2, 4, 6, 8\}$ and $B := \{1, 2, 3, 4, 5, 6\}$. Then $A \cup B = \{1, 2, 3, 4, 5, 6, 8\}$ and $A \cap B = \{2, 4, 6\}$. Also, we write that 1 is an element of B like this: $1 \in B$. Of course, 1 is not an element of A, so we write $1 \notin A$. If we have another set $C := \{4, 8\}$, and we notice that every element of C is an element of A, we say that C is a subset of A, which is written $C \subset A$. We can see that there are elements in A which are not in C. We can describe these as $A \setminus C = \{2, 6\}$.

These operations can be indexed. Suppose that $A_1, A_2, \ldots, A_m$ are m sets. We can write an indexed union or intersection as follows:

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

Similarly, if we have a sequence of numbers, $a_1, a_2, \ldots, a_m$, we can compute their indexed sum as follows:

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

If the context is clear, this may be abbreviated as

$$\sum_{i} a_i.$$

We use the binomial coefficient $\binom{n}{k}$, which means

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

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

Here, and throughout the book, $X \lesssim Y$ means that as X and Y grow large, typically as a function of some parameter, say N, there exists a positive constant C, which does not depend on N, such that $X \leq CY$. This is also sometimes written $X = O(Y)$, and is read X is big "O" of Y, or on the order of Y. Furthermore, $X \approx Y$ means that $X \lesssim Y$ and $Y \lesssim X$. We take this notational game a step further and write $X \lessapprox Y$ if for every $\epsilon > 0$ there exists $C_\epsilon > 0$ such that $X \leq C_\epsilon N^\epsilon Y$. For example, $N \log^{100}(N) \lessapprox N$. This notation is not only more convenient, but it also emphasizes the fact that these constants do not affect our results asymptotically.

Naturally, as the the theory develops, we will use more symbols and shorthand, but these will all be introduced as they arise. Also, when we define anything new, we will *italicize* the new term.

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

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

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

and

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

Exercises

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

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

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

and no greater than $dn^{\frac{2}{d}}$. Now check that

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

follows from this observation.

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

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

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

Chapter 1

The $\sqrt{n}$ theory

1. Erdős' original argument

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

Theorem 1.1 (Erdős [**12**]). *Suppose that $d = 2$ and $\#P = n$. Then $\#\Delta(P) \gtrsim n^{\frac{1}{2}}$.*

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

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

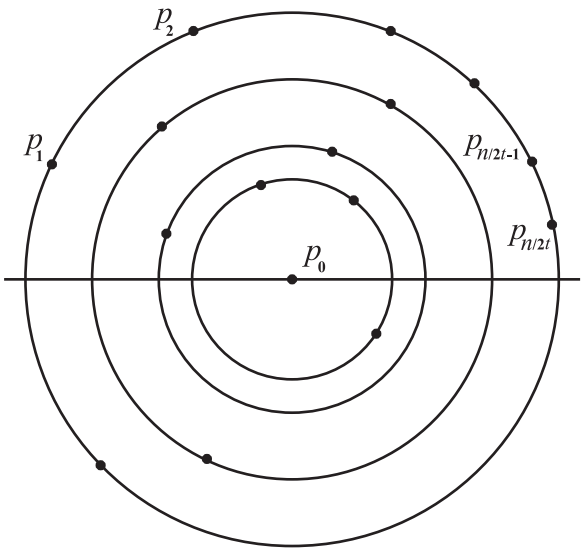

Figure 1.1. Circles about p_0 and the East-West line.

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

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

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

There are several ways to proceed here. One way is to "guess" the answer. Since we already took care of the case where $t \geq \sqrt{n}$, we

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

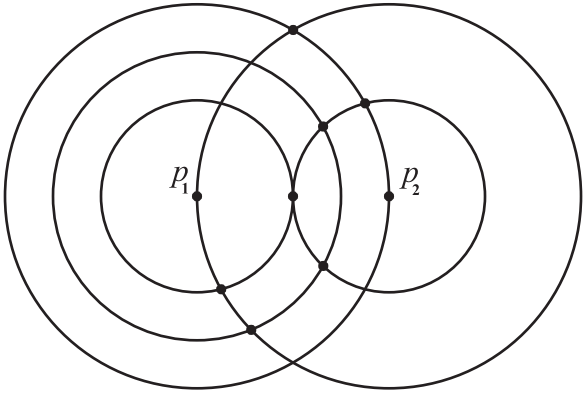

Figure 1.2. Circles about p_1 and p_2 that cover P.

can assume that $t < \sqrt{n}$. Then $n/2t > \sqrt{n}/2$, so either way,

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

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

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

This transforms (1.1) into (1.2). □

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

2. Higher dimensions

What about higher dimensions? We try the same approach. Choose a point in P and draw all spheres that contain at least one point of P. As before, let t denote the number of these spheres. If t is large enough, we are done. If not, then one of the spheres contains at least n/t points. Unfortunately, if $d > 2$, we cannot run the simple-minded argument that worked in two dimensions. Or can we? Notice that if

we are working in $\mathbb{R}^d$, the surface of each sphere is $(d-1)$-dimensional, whatever that means. This suggests the following approach, which uses induction. If you are unfamiliar with proofs by induction, Appendix C has a brief explanation of this concept.

Let S^k denote the k-dimensional sphere. So S^1 is the circle, S^2 would be a hollow spherical shell, like a basketball, and so on.

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

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

In the case of $\mathbb{R}^2$, the induction hypothesis holds by Theorem 1.1. Similarly, we have verified the statement for S^1 in the proof of Theorem 1.1, when we noticed that if there were a number of points on one of the circles, then they must determine about that many distinct distances. We are now ready to complete the argument for higher dimensions. When we follow this reasoning in dimension d, we end up with t $(d-1)$-spheres, one of which must have at least n/t points on it as in the $d = 2$ proof. By induction, these points determine $\gtrsim \left(\frac{n}{t}\right)^{\frac{1}{d-1}}$ distances. It follows that

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

We now use the fact that

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

which implies that

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

We have just proved the following result.

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

Most of our focus will be on the the problem in the plane; however, there has been a fair amount of work done in higher dimensions. In [**4**], a bound of $n^{77/141-\epsilon}$, for any $\epsilon > 0$, is achieved for three dimensions. In [**47**], a general lower bound of $n^{2/d-2/(d(d+1))}$ is attained for $d \geq 4$, improving the earlier work in [**46**].

3. Arbitrary metrics

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

We now present a formal definition of a general metric.

Definition 1.1. We call a function, $d(x, y)$, on a set, S, a *metric* if it returns a real number for any two elements of S satisfying the following for all distinct $x, y, z \in S$:

(i) $d(x, x) = 0$;

(ii) $d(x, y) > 0$;

(iii) $d(x, y) = d(y, x)$ (symmetry);

(iv) $d(x, z) \geq d(x, y) + d(y, z)$ (triangle inequality).

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

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

It is customary to think of the distance from one point to another as the length of the straight line connecting the two points. However, as our cursory exploration of the taxicab metric suggests, this does

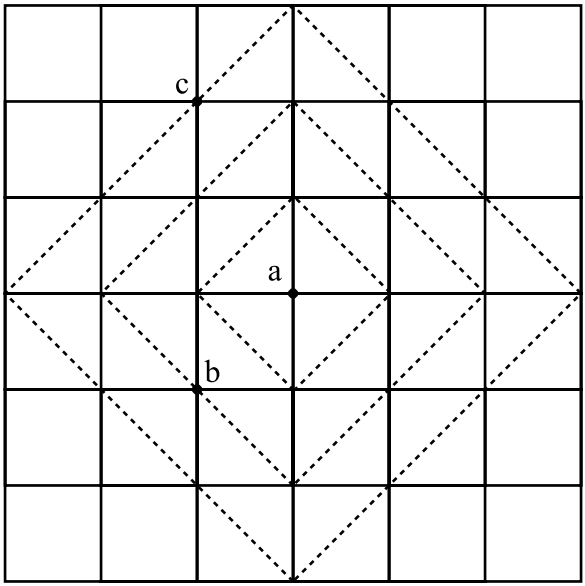

Figure 1.3. The grid represents an overhead view of a city. If you are located at a, you will have to walk two blocks to b, or three blocks to c. The dashed lines represent three dilates of the l_1 circle.

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

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

In the arguments above, not all of the properties of the standard Euclidean circle were utilized. Exercises 1.6 and 1.7 accentuate some

of the critical similarities and differences between arbitrary metrics and the Euclidean metric.

At this point, we could spend a long time introducing and developing many different types of metrics, but instead, we want you to discover on your own what types of objects can be viewed as metrics, and in what sense. As you read through this book, other types of metrics and metric-like objects will naturally come along. In mathematics, it is rare that a definition magically descends from the sky and dares us to explore its uses. Typically, various scenarios give rise to sensible constraints on a useful object, which are then compiled into a definition sometime after the subject has been investigated a little. For this book in particular, we feel that it is far more instructive to watch the theory grow by necessity than to introduce a laundry list of definitions and then draw conclusions. If you can come up with some of your own variations on the examples given in Exercise 1.8, you will get more out of this book.

Exercises

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

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

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

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

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

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

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

Exercise 1.8. Consider the following metric-like objects. Assume that they all map $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, or that they take two points in the plane as input and give one number as output. Determine which are genuine metrics, and which are not. Could one sensibly ask questions like the Erdős distance problem of these objects? If $x = (x_1, x_2)$ and $y = (y_1, y_2)$, then

(1) $F(x, y) = |x| + |y|$;
(2) $D(x, y) = x_1x_2 + y_1y_2$;
(3) $\Phi(x, y) = \frac{|x-y|}{|x+y|+1}$.

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

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

Exercise 1.10. We have been considering how many different distances are determined by a point set. Another question is to ask how often a single distance can occur. This is referred to as the *unit distance problem*. Why do we only need to consider unit[3] distances? Consider $\mathbb{R}^4$, and call the coordinate axes x, y, z, and w. Arrange $\frac{n}{2}$ points in a circle of radius $\frac{\sqrt{2}}{2}$ in the plane determined by the x and y axes, centered at the origin. Then arrange $\frac{n}{2}$ points in another circle of radius $\frac{\sqrt{2}}{2}$ in the plane determined by the z and w axes. How often does the unit distance occur? This is called a *Lenz construction*.

[3]Here, *unit distance* means a distance equal to one.