nation based on regurgitating extracts from the book. (Indeed, in my own examinations I gave a supplemental sheet listing the key definitions and theorems which were relevant to the examination problems.) Making the examinations similar to the homework assigned in the course will also help motivate the students to work through and understand their homework problems as thoroughly as possible (as opposed to, say, using flash cards or other such devices to memorize material), which is good preparation not only for examinations but for doing mathematics in general.

Some of the material in this textbook is somewhat peripheral to the main theme and may be omitted for reasons of time constraints. For instance, as set theory is not as fundamental to analysis as are the number systems, the chapters on set theory (Chapters 3, 8) can be covered more quickly and with substantially less rigour, or be given as reading assignments. The appendices on logic and the decimal system are intended as optional or supplemental reading and would probably not be covered in the main course lectures; the appendix on logic is particularly suitable for reading concurrently with the first few chapters. Also, Chapter 16 (on Fourier series) is not needed elsewhere in the text and can be omitted.

For reasons of length, this textbook has been split into two volumes. The first volume is slightly longer, but can be covered in about thirty lectures if the peripheral material is omitted or abridged. The second volume refers at times to the first, but can also be taught to students who have had a first course in analysis from other sources. It also takes about thirty lectures to cover.

I am deeply indebted to my students, who over the progression of the real analysis course corrected several errors in the lectures notes from which this text is derived, and gave other valuable feedback. I am also very grateful to the many anonymous referees who made several corrections and suggested many important improvements to the text.

Terence Tao

Chapter 12

Metric spaces

12.1 Definitions and examples

In Definition 6.1.5 we defined what it meant for a sequence $(x_n)_{n=m}^{\infty}$ of real numbers to converge to another real number x; indeed, this meant that for every $\varepsilon > 0$, there exists an $N \geq m$ such that $|x - x_n| \leq \varepsilon$ for all $n \geq N$. When this is the case, we write $\lim_{n\to\infty} x_n = x$.

Intuitively, when a sequence $(x_n)_{n=m}^{\infty}$ converges to a limit x, this means that somehow the elements x_n of that sequence will eventually be as close to x as one pleases. One way to phrase this more precisely is to introduce the *distance function* $d(x, y)$ between two real numbers by $d(x, y) := |x-y|$. (Thus for instance $d(3, 5) = 2$, $d(5, 3) = 2$, and $d(3, 3) = 0$.) Then we have

Lemma 12.1.1. *Let $(x_n)_{n=m}^{\infty}$ be a sequence of real numbers, and let x be another real number. Then $(x_n)_{n=m}^{\infty}$ converges to x if and only if* $\lim_{n\to\infty} d(x_n, x) = 0$.

Proof. See Exercise 12.1.1. $\square$

One would now like to generalize this notion of convergence, so that one can take limits not just of sequences of real numbers, but also sequences of complex numbers, or sequences of vectors, or sequences of matrices, or sequences of functions, even sequences of sequences. One way to do this is to redefine the notion of convergence each time we deal with a new type of object. As you can guess, this will quickly get tedious. A more efficient way is to work *abstractly*, defining a very general class of spaces - which includes such standard spaces as the real numbers, complex numbers, vectors, etc. - and define the notion of convergence on this entire class of spaces at once. (A *space* is just the set of all objects of a certain type - the space of all real numbers, the space of all 3×3 matrices, etc. Mathematically, there is not much distinction between a space and a set, except that spaces tend to have much more structure than what a random set would have. For instance, the space of real numbers comes with operations such as addition and multiplication, while a general set would not.)

It turns out that there are two very useful classes of spaces which do the job. The first class is that of *metric spaces*, which we will study here. There is a more general class of spaces, called *topological spaces*, which is also very important, but we will only deal with this generalization briefly, in Section 13.5.

Roughly speaking, a metric space is any space X which has a concept of *distance* $d(x, y)$ - and this distance should behave in a reasonable manner. More precisely, we have

Definition 12.1.2 (Metric spaces). A *metric space* (X, d) is a space X of objects (called *points*), together with a *distance function* or *metric* $d : X \times X \to [0, +\infty)$, which associates to each pair x, y of points in X a non-negative real number $d(x, y) \geq 0$. Furthermore, the metric must satisfy the following four axioms:

(a) For any $x \in X$, we have $d(x, x) = 0$.

(b) (Positivity) For any *distinct* $x, y \in X$, we have $d(x, y) > 0$.

(c) (Symmetry) For any $x, y \in X$, we have $d(x, y) = d(y, x)$.

(d) (Triangle inequality) For any $x, y, z \in X$, we have $d(x, z) \leq d(x, y) + d(y, z)$.

In many cases it will be clear what the metric d is, and we shall abbreviate (X, d) as just X.

Remark 12.1.3. The conditions (a) and (b) can be rephrased as follows: for any $x, y \in X$ we have $d(x, y) = 0$ if and only if $x = y$. (Why is this equivalent to (a) and (b)?)

Example 12.1.4 (The real line). Let $\mathbf{R}$ be the real numbers, and let $d : \mathbf{R} \times \mathbf{R} \to [0, \infty)$ be the metric $d(x, y) := |x - y|$ mentioned earlier. Then $(\mathbf{R}, d)$ is a metric space (Exercise 12.1.2). We refer to d as the *standard metric* on $\mathbf{R}$, and if we refer to $\mathbf{R}$ as a metric space, we assume that the metric is given by the standard metric d unless otherwise specified.

Example 12.1.5 (Induced metric spaces). Let (X, d) be any metric space, and let Y be a subset of X. Then we can restrict the metric function $d : X \times X \to [0, +\infty)$ to the subset $Y \times Y$ of $X \times X$ to create a restricted metric function $d|_{Y \times Y} : Y \times Y \to [0, +\infty)$ of Y; this is known as the metric on Y *induced* by the metric d on X. The pair $(Y, d|_{Y \times Y})$ is a metric space (Exercise 12.1.4) and is known the *subspace* of (X, d) induced by Y. Thus for instance the metric on the real line in the previous example induces a metric space structure on any subset of the reals, such as the integers $\mathbf{Z}$, or an interval $[a, b]$, etc.

Example 12.1.6 (Euclidean spaces). Let $n \geq 1$ be a natural number, and let $\mathbf{R}^n$ be the space of n-tuples of real numbers:

$$\mathbf{R}^n = \{(x_1, x_2, \ldots, x_n) : x_1, \ldots, x_n \in \mathbf{R}\}.$$

We define the *Euclidean metric* (also called the *l^2 metric*) $d_{l^2} : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ by

$$\begin{aligned} d_{l^2}((x_1, \ldots, x_n), (y_1, \ldots, y_n)) &:= \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2} \\ &= (\sum_{i=1}^{n} (x_i - y_i)^2)^{1/2}. \end{aligned}$$

Thus for instance, if $n = 2$, then $d_{l^2}((1, 6), (4, 2)) = \sqrt{3^2 + 4^2} = 5$. This metric corresponds to the geometric distance between the two points $(x_1, x_2, \ldots, x_n)$, $(y_1, y_2, \ldots, y_n)$ as given by Pythagoras' theorem. (We remark however that while geometry does give some very important examples of metric spaces, it is possible to have metric spaces which have no obvious geometry whatsoever. Some examples are given below.) The verification that $(\mathbf{R}^n, d)$ is indeed a metric space can be seen geometrically (for instance, the triangle inequality now asserts that the length of one side of a triangle is always less than or equal to the sum of the lengths of the other two sides), but can also be proven algebraically (see Exercise 12.1.6). We refer to $(\mathbf{R}^n, d_{l^2})$ as the *Euclidean space* of *dimension* n.

Example 12.1.7 (Taxi-cab metric). Again let $n \geq 1$, and let $\mathbf{R}^n$ be as before. But now we use a different metric d_{l^1}, the so-called *taxicab metric* (or l^1 *metric*), defined by

$$\begin{aligned} d_{l^1}((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)) &:= |x_1 - y_1| + \ldots + |x_n - y_n| \\ &= \sum_{i=1}^{n} |x_i - y_i|. \end{aligned}$$

Thus for instance, if $n = 2$, then $d_{l^1}((1,6),(4,2)) = 5 + 2 = 7$. This metric is called the taxi-cab metric, because it models the distance a taxi-cab would have to traverse to get from one point to another if the cab was only allowed to move in cardinal directions (north, south, east, west) and not diagonally. As such it is always at least as large as the Euclidean metric, which measures distance "as the crow flies", as it were. We claim that the space $(\mathbf{R}^n, d_{l^1})$ is also a metric space (Exercise 12.1.7). The metrics are not quite the same, but we do have the inequalities

$$d_{l^2}(x, y) \leq d_{l^1}(x, y) \leq \sqrt{n} d_{l^2}(x, y) \qquad (12.1)$$

for all x, y (see Exercise 12.1.8).

Remark 12.1.8. The taxi-cab metric is useful in several places, for instance in the theory of error correcting codes. A string of n binary digits can be thought of as an element of $\mathbf{R}^n$, for instance the binary string 10010 can be thought of as the point $(1, 0, 0, 1, 0)$ in $\mathbf{R}^5$. The taxi-cab distance between two binary strings is then the number of bits in the two strings which do not match, for instance $d_{l^1}(10010, 10101) = 3$. The goal of error-correcting codes is to encode each piece of information (e.g., a letter of the alphabet) as a binary string in such a way that all the binary strings are as far away in the taxicab metric from each other as possible; this minimizes the chance that any distortion of the bits due to random noise can accidentally change one of the coded binary strings to another, and also maximizes the chance that any such distortion can be detected and correctly repaired.

Example 12.1.9 (Sup norm metric). Again let $n \geq 1$, and let $\mathbf{R}^n$ be as before. But now we use a different metric d_{l^∞}, the so-called *sup norm metric* (or *l^∞ metric*), defined by

$$d_{l^\infty}((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)) := \sup\{|x_i - y_i| : 1 \leq i \leq n\}.$$

Thus for instance, if $n = 2$, then $d_{l^\infty}((1,6),(4,2)) = \sup(5,2) = 7$. The space $(\mathbf{R}^n, d_{l^\infty})$ is also a metric space (Exercise 12.1.9), and is related to the l^2 metric by the inequalities

$$\frac{1}{\sqrt{n}} d_{l^2}(x,y) \leq d_{l^\infty}(x,y) \leq d_{l^2}(x,y) \tag{12.2}$$

for all x, y (see Exercise 12.1.10).

Remark 12.1.10. The l^1, l^2, and l^∞ metrics are special cases of the more general *l^p metrics*, where $p \in [1, +\infty]$, but we will not discuss these more general metrics in this text.

Example 12.1.11 (Discrete metric). Let X be an arbitrary set (finite or infinite), and define the *discrete metric d_{disc}* by setting $d_{disc}(x,y) := 0$ when $x = y$, and $d_{disc}(x,y) := 1$ when $x \neq y$. Thus, in this metric, all points are equally far apart. The space (X, d_{disc}) is a metric space (Exercise 12.1.11). Thus every set X has at least one metric on it.

Example 12.1.12 (Geodesics). (Informal) Let X be the sphere $\{(x,y,z) \in \mathbf{R}^3 : x^2 + y^2 + z^2 = 1\}$, and let $d((x,y,z),(x',y',z'))$ be the length of the shortest curve in X which starts at (x,y,z) and ends at (x',y',z'). (This curve turns out to be an arc of a great circle; we will not prove this here, as it requires *calculus of variations*, which is beyond the scope of this text.) This makes X into a metric space; the reader should be able to verify (without using any geometry of the sphere) that the triangle inequality is more or less automatic from the definition.

Example 12.1.13 (Shortest paths). (Informal) Examples of metric spaces occur all the time in real life. For instance, X could be all the computers currently connected to the internet, and $d(x, y)$ is the shortest number of connections it would take for a packet to travel from computer x to computer y; for instance, if x and y are not directly connected, but are both connected to z, then $d(x, y) = 2$. Assuming that all computers in the internet can ultimately be connected to all other computers (so that $d(x, y)$ is always finite), then (X, d) is a metric space (why?). Games such as "six degrees of separation" are also taking place in a similar metric space (what is the space, and what is the metric, in this case?). Or, X could be a major city, and $d(x, y)$ could be the shortest time it takes to drive from x to y (although this space might not satisfy axiom (iii) in real life!).

Now that we have metric spaces, we can define convergence in these spaces.

Definition 12.1.14 (Convergence of sequences in metric spaces). Let m be an integer, (X, d) be a metric space and let $(x^{(n)})_{n=m}^{\infty}$ be a sequence of points in X (i.e., for every natural number $n \geq m$, we assume that $x^{(n)}$ is an element of X). Let x be a point in X. We say that $(x^{(n)})_{n=m}^{\infty}$ *converges to x with respect to the metric d*, if and only if the limit $\lim_{n\to\infty} d(x^{(n)}, x)$ exists and is equal to 0. In other words, $(x^{(n)})_{n=m}^{\infty}$ converges to x with respect to d if and only if for every $\varepsilon > 0$, there exists an $N \geq m$ such that $d(x^{(n)}, x) \leq \varepsilon$ for all $n \geq N$. (Why are these two definitions equivalent?)

Remark 12.1.15. In view of Lemma 12.1.1 we see that this definition generalizes our existing notion of convergence of sequences of real numbers. In many cases, it is obvious what the metric d

is, and so we shall often just say "$(x^{(n)})_{n=m}^{\infty}$ converges to x" instead of "$(x^{(n)})_{n=m}^{\infty}$ converges to x with respect to the metric d" when there is no chance of confusion. We also sometimes write "$x^{(n)} \to x$ as $n \to \infty$" instead.

Remark 12.1.16. There is nothing special about the superscript n in the above definition; it is a dummy variable. Saying that $(x^{(n)})_{n=m}^{\infty}$ converges to x is exactly the same statement as saying that $(x^{(k)})_{k=m}^{\infty}$ converges to x, for example; and sometimes it is convenient to change superscripts, for instance if the variable n is already being used for some other purpose. Similarly, it is not necessary for the sequence $x^{(n)}$ to be denoted using the superscript (n); the above definition is also valid for sequences x_n, or functions $f(n)$, or indeed of any expression which depends on n and takes values in X. Finally, from Exercises 6.1.3, 6.1.4 we see that the starting point m of the sequence is unimportant for the purposes of taking limits; if $(x^{(n)})_{n=m}^{\infty}$ converges to x, then $(x^{(n)})_{n=m'}^{\infty}$ also converges to x for any $m' \geq m$.

Example 12.1.17. We work in the Euclidean space $\mathbf{R}^2$ with the standard Euclidean metric d_{l^2}. Let $(x^{(n)})_{n=1}^{\infty}$ denote the sequence $x^{(n)} := (1/n, 1/n)$ in $\mathbf{R}^2$, i.e., we are considering the sequence $(1,1), (1/2,1/2), (1/3,1/3), \ldots$. Then this sequence converges to $(0,0)$ with respect to the Euclidean metric d_{l^2}, since

$$\lim_{n\to\infty} d_{l^2}(x^{(n)}, (0,0)) = \lim_{n\to\infty} \sqrt{\frac{1}{n^2} + \frac{1}{n^2}} = \lim_{n\to\infty} \frac{\sqrt{2}}{n} = 0.$$

The sequence $(x^{(n)})_{n=1}^{\infty}$ also converges to $(0,0)$ with respect to the taxi-cab metric d_{l^1}, since

$$\lim_{n\to\infty} d_{l^1}(x^{(n)}, (0,0)) = \lim_{n\to\infty} \frac{1}{n} + \frac{1}{n} = \lim_{n\to\infty} \frac{2}{n} = 0.$$

Similarly the sequence converges to $(0,0)$ in the sup norm metric d_{l^∞} (why?). However, the sequence $(x^{(n)})_{n=1}^{\infty}$ does *not* converge to $(0,0)$ in the discrete metric d_{disc}, since

$$\lim_{n\to\infty} d_{disc}(x^{(n)}, (0,0)) = \lim_{n\to\infty} 1 = 1 \neq 0.$$

Thus the convergence of a sequence can depend on what metric one uses[1].

In the case of the above four metrics - Euclidean, taxi-cab, sup norm, and discrete - it is in fact rather easy to test for convergence.

Proposition 12.1.18 (Equivalence of l^1, l^2, l^∞). *Let $\mathbf{R}^n$ be a Euclidean space, and let $(x^{(k)})_{k=m}^\infty$ be a sequence of points in $\mathbf{R}^n$. We write $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})$, i.e., for $j = 1, 2, \ldots, n$, $x_j^{(k)} \in \mathbf{R}$ is the j^{th} co-ordinate of $x^{(k)} \in \mathbf{R}^n$. Let $x = (x_1, \ldots, x_n)$ be a point in $\mathbf{R}^n$. Then the following four statements are equivalent:*

(a) *$(x^{(k)})_{k=m}^\infty$ converges to x with respect to the Euclidean metric d_{l^2}.*

(b) *$(x^{(k)})_{k=m}^\infty$ converges to x with respect to the taxi-cab metric d_{l^1}.*

(c) *$(x^{(k)})_{k=m}^\infty$ converges to x with respect to the sup norm metric d_{l^∞}.*

(d) *For every $1 \leq j \leq n$, the sequence $(x_j^{(k)})_{k=m}^\infty$ converges to x_j. (Notice that this is a sequence of real numbers, not of points in $\mathbf{R}^n$.)*

Proof. See Exercise 12.1.12. □

In other words, a sequence converges in the Euclidean, taxi-cab, or sup norm metric if and only if each of its components converges individually. Because of the equivalence of (a), (b) and (c), we say that the Euclidean, taxicab, and sup norm metrics on $\mathbf{R}^n$ are *equivalent*. (There are infinite-dimensional analogues

[1]For a somewhat whimsical real-life example, one can give a city an "automobile metric", with $d(x, y)$ defined as the time it takes for a car to drive from x to y, or a "pedestrian metric", where $d(x, y)$ is the time it takes to walk on foot from x to y. (Let us assume for sake of argument that these metrics are symmetric, though this is not always the case in real life.) One can easily imagine examples where two points are close in one metric but not another.

of the Euclidean, taxicab, and sup norm metrics which are *not* equivalent, see for instance Exercise 12.1.15.)

For the discrete metric, convergence is much rarer: the sequence must be eventually constant in order to converge.

Proposition 12.1.19 (Convergence in the discrete metric). *Let X be any set, and let d_{disc} be the discrete metric on X. Let $(x^{(n)})_{n=m}^{\infty}$ be a sequence of points in X, and let x be a point in X. Then $(x^{(n)})_{n=m}^{\infty}$ converges to x with respect to the discrete metric d_{disc} if and only if there exists an $N \geq m$ such that $x^{(n)} = x$ for all $n \geq N$.*

Proof. See Exercise 12.1.13. □

We now prove a basic fact about converging sequences; they can only converge to at most one point at a time.

Proposition 12.1.20 (Uniqueness of limits). *Let (X, d) be a metric space, and let $(x^{(n)})_{n=m}^{\infty}$ be a sequence in X. Suppose that there are two points $x, x' \in X$ such that $(x^{(n)})_{n=m}^{\infty}$ converges to x with respect to d, and $(x^{(n)})_{n=m}^{\infty}$ also converges to x' with respect to d. Then we have $x = x'$.*

Proof. See Exercise 12.1.14. □

Because of the above Proposition, it is safe to introduce the following notation: if $(x^{(n)})_{n=m}^{\infty}$ converges to x in the metric d, then we write $d - \lim_{n\to\infty} x^{(n)} = x$, or simply $\lim_{n\to\infty} x^{(n)} = x$ when there is no confusion as to what d is. For instance, in the example of $(\frac{1}{n}, \frac{1}{n})$, we have

$$d_{l^2} - \lim_{n\to\infty} (\frac{1}{n}, \frac{1}{n}) = d_{l^1} - \lim_{n\to\infty} (\frac{1}{n}, \frac{1}{n}) = (0, 0),$$

but $d_{disc} - \lim_{n\to\infty}(\frac{1}{n}, \frac{1}{n})$ is undefined. Thus the meaning of $d - \lim_{n\to\infty} x^{(n)}$ can depend on what d is; however Proposition 12.1.20 assures us that once d is fixed, there can be at most one value of $d - \lim_{n\to\infty} x^{(n)}$. (Of course, it is still possible that this limit does not exist; some sequences are not convergent.) Note that by Lemma 12.1.1, this definition of limit generalizes the notion of limit in Definition 6.1.8.

Remark 12.1.21. It is possible for a sequence to converge to one point using one metric, and another point using a different metric, although such examples are usually quite artificial. For instance, let $X := [0, 1]$, the closed interval from 0 to 1. Using the usual metric d, we have $d - \lim_{n\to\infty} \frac{1}{n} = 0$. But now suppose we "swap" the points 0 and 1 in the following manner. Let $f : [0, 1] \to [0, 1]$ be the function defined by $f(0) := 1$, $f(1) := 0$, and $f(x) := x$ for all $x \in (0, 1)$, and then define $d'(x, y) := d(f(x), f(y))$. Then (X, d') is still a metric space (why?), but now $d' - \lim_{n\to\infty} \frac{1}{n} = 1$. Thus changing the metric on a space can greatly affect the nature of convergence (also called the *topology*) on that space; see Section 13.5 for a further discussion of topology.

Exercise 12.1.1. Prove Lemma 12.1.1.

Exercise 12.1.2. Show that the real line with the metric $d(x, y) := |x-y|$ is indeed a metric space. (Hint: you may wish to review your proof of Proposition 4.3.3.)

Exercise 12.1.3. Let X be a set, and let $d : X \times X \to [0, \infty)$ be a function.

(a) Give an example of a pair (X, d) which obeys axioms (bcd) of Definition 12.1.2, but not (a). (Hint: modify the discrete metric.)

(b) Give an example of a pair (X, d) which obeys axioms (acd) of Definition 12.1.2, but not (b).

(c) Give an example of a pair (X, d) which obeys axioms (abd) of Definition 12.1.2, but not (c).

(d) Give an example of a pair (X, d) which obeys axioms (abc) of Definition 12.1.2, but not (d). (Hint: try examples where X is a finite set.)

Exercise 12.1.4. Show that the pair $(Y, d|_{Y\times Y})$ defined in Example 12.1.5 is indeed a metric space.

Exercise 12.1.5. Let $n \geq 1$, and let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ be real numbers. Verify the identity

$$(\sum_{i=1}^{n} a_i b_i)^2 + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i b_j - a_j b_i)^2 = (\sum_{i=1}^{n} a_i^2)(\sum_{j=1}^{n} b_j^2),$$

and conclude the *Cauchy-Schwarz inequality*

$$|\sum_{i=1}^{n} a_i b_i| \leq (\sum_{i=1}^{n} a_i^2)^{1/2} (\sum_{j=1}^{n} b_j^2)^{1/2}. \quad (12.3)$$

Then use the Cauchy-Schwarz inequality to prove the *triangle inequality*

$$(\sum_{i=1}^{n} (a_i + b_i)^2)^{1/2} \leq (\sum_{i=1}^{n} a_i^2)^{1/2} + (\sum_{j=1}^{n} b_j^2)^{1/2}.$$

Exercise 12.1.6. Show that $(\mathbf{R}^n, d_{l^2})$ in Example 12.1.6 is indeed a metric space. (Hint: use Exercise 12.1.5.)

Exercise 12.1.7. Show that the pair $(\mathbf{R}^n, d_{l^1})$ in Example 12.1.7 is indeed a metric space.

Exercise 12.1.8. Prove the two inequalities in (12.1). (For the first inequality, square both sides. For the second inequality, use Exercise (12.1.5).

Exercise 12.1.9. Show that the pair $(\mathbf{R}^n, d_{l^\infty})$ in Example 12.1.9 is indeed a metric space.

Exercise 12.1.10. Prove the two inequalities in (12.2).

Exercise 12.1.11. Show that the discrete metric $(\mathbf{R}^n, d_{disc})$ in Example 12.1.11 is indeed a metric space.

Exercise 12.1.12. Prove Proposition 12.1.18.

Exercise 12.1.13. Prove Proposition 12.1.19.

Exercise 12.1.14. Prove Proposition 12.1.20. (Hint: modify the proof of Proposition 6.1.7.)

Exercise 12.1.15. Let

$$X := \{(a_n)_{n=0}^{\infty} : \sum_{n=0}^{\infty} |a_n| < \infty\}$$

be the space of absolutely convergent sequences. Define the l^1 and l^∞ metrics on this space by

$$d_{l^1}((a_n)_{n=0}^{\infty}, (b_n)_{n=0}^{\infty}) := \sum_{n=0}^{\infty} |a_n - b_n|;$$

$$d_{l^\infty}((a_n)_{n=0}^{\infty}, (b_n)_{n=0}^{\infty}) := \sup_{n \in \mathbf{N}} |a_n - b_n|.$$

Show that these are both metrics on X, but show that there exist sequences $x^{(1)}, x^{(2)}, \ldots$ of elements of X (i.e., sequences of sequences) which are convergent with respect to the d_{l^∞} metric but not with respect to the d_{l^1} metric. Conversely, show that any sequence which converges in the d_{l^1} metric automatically converges in the d_{l^∞} metric.

Exercise 12.1.16. Let $(x_n)_{n=1}^\infty$ and $(y_n)_{n=1}^\infty$ be two sequences in a metric space (X, d). Suppose that $(x_n)_{n=1}^\infty$ converges to a point $x \in X$, and $(y_n)_{n=1}^\infty$ converges to a point $y \in X$. Show that $\lim_{n\to\infty} d(x_n, y_n) = d(x, y)$. (Hint: use the triangle inequality several times.)

12.2 Some point-set topology of metric spaces

Having defined the operation of convergence on metric spaces, we now define a couple other related notions, including that of open set, closed set, interior, exterior, boundary, and adherent point. The study of such notions is known as *point-set topology*, which we shall return to in Section 13.5.

We first need the notion of a *metric ball*, or more simply a *ball*.

Definition 12.2.1 (Balls). Let (X, d) be a metric space, let x_0 be a point in X, and let $r > 0$. We define the *ball* $B_{(X,d)}(x_0, r)$ in X, centered at x_0, and with radius r, in the metric d, to be the set

$$B_{(X,d)}(x_0, r) := \{x \in X : d(x, x_0) < r\}.$$

When it is clear what the metric space (X, d) is, we shall abbreviate $B_{(X,d)}(x_0, r)$ as just $B(x_0, r)$.

Example 12.2.2. In $\mathbf{R}^2$ with the Euclidean metric d_{l^2}, the ball $B_{(\mathbf{R}^2, d_{l^2})}((0,0), 1)$ is the open disc

$$B_{(\mathbf{R}^2, d_{l^2})}((0,0), 1) = \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 < 1\}.$$

However, if one uses the taxi-cab metric d_{l^1} instead, then we obtain a diamond:

$$B_{(\mathbf{R}^2, d_{l^1})}((0,0), 1) = \{(x, y) \in \mathbf{R}^2 : |x| + |y| < 1\}.$$

If we use the discrete metric, the ball is now reduced to a single point:

$$B_{(\mathbf{R}^2, d_{disc})}((0,0),1) = \{(0,0)\},$$

although if one increases the radius to be larger than 1, then the ball now encompasses all of $\mathbf{R}^2$. (Why?)

Example 12.2.3. In $\mathbf{R}$ with the usual metric d, the open interval $(3,7)$ is also the metric ball $B_{(\mathbf{R},d)}(5,2)$.

Remark 12.2.4. Note that the smaller the radius r, the smaller the ball $B(x_0, r)$. However, $B(x_0, r)$ always contains at least one point, namely the center x_0, as long as r stays positive, thanks to Definition 12.1.2(a). (We don't consider balls of zero radius or negative radius since they are rather boring, being just the empty set.)

Using metric balls, one can now take a set E in a metric space X, and classify three types of points in X: interior, exterior, and boundary points of E.

Definition 12.2.5 (Interior, exterior, boundary). Let (X, d) be a metric space, let E be a subset of X, and let x_0 be a point in X. We say that x_0 is an *interior point of* E if there exists a radius $r > 0$ such that $B(x_0, r) \subseteq E$. We say that x_0 is an *exterior point of* E if there exists a radius $r > 0$ such that $B(x_0, r) \cap E = \emptyset$. We say that x_0 is a *boundary point of* E if it is neither an interior point nor an exterior point of E.

The set of all interior points of E is called the *interior* of E and is sometimes denoted $\mathrm{int}(E)$. The set of exterior points of E is called the *exterior* of E and is sometimes denoted $\mathrm{ext}(E)$. The set of boundary points of E is called the *boundary* of E and is sometimes denoted ∂E.

Remark 12.2.6. If x_0 is an interior point of E, then x_0 must actually be an element of E, since balls $B(x_0, r)$ always contain their center x_0. Conversely, if x_0 is an exterior point of E, then x_0 cannot be an element of E. In particular it is not possible for