Metric Spaces and Point-Set Topology

Nicolas Addington

Last updated October 18, 2024

These notes are a slightly expanded version of what I will say in lecture. I will update them throughout the quarter.

Contents

1 Continuity, Convergence, and Metric Spaces

In a previous course you may have seen the definition of continuity and convergence in $\mathbb{R}^n$:

Definition 1.1.

(a) A map $f\colon \mathbb{R}^m \to \mathbb{R}^n$ is *continuous* at a point $\mathbf{a} \in \mathbb{R}^m$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that for all $\mathbf{x} \in \mathbb{R}^m$ with $|\mathbf{x} - \mathbf{a}| < \delta$ we have $|f(\mathbf{x}) - f(\mathbf{a})| < \epsilon$.

(b) A sequence of points $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots \in \mathbb{R}^n$ *converges* to a limit $\ell \in \mathbb{R}^n$ if for every $\epsilon > 0$ there is a natural number N such that for all $n \geq N$ we have $|\mathbf{x}_n - \ell| < \epsilon$.

You don't have to be completely at ease with these definitions, but if you haven't at least worked with them in one variable – that is, with functions $f\colon \mathbb{R} \to \mathbb{R}$ and sequences of numbers $x_1, x_2, x_3, \ldots \in \mathbb{R}$ – then you should take a rigorous advanced calculus course, called introductory real analysis in some places, before taking this one. My favorite book is Spivak's *Calculus*. Oregon's Math 316–317 currently uses Abbott's *Understanding Analysis*.

In the definition above, $|\mathbf{x} - \mathbf{a}|$ means the length of the vector $\mathbf{x} - \mathbf{a}$, which we should understand as measuring the distance between $\mathbf{x}$ and $\mathbf{a}$ in $\mathbb{R}^m$, and similarly with $|f(\mathbf{x}) - f(\mathbf{a})|$ and $|\mathbf{x}_n - \ell|$. This course begins from the observation that when we study continuous maps and convergent sequences, we can forget almost everything we know about vectors and their geometry, and retain only the notion of distance. We introduce the following definition.

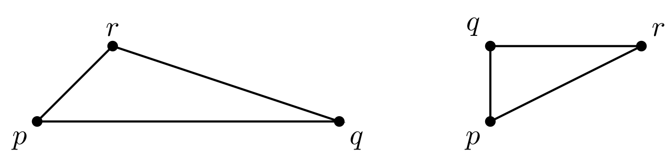
Definition 1.2. A *metric* on a set X is a function $d\colon X \times X \to \mathbb{R}_{\geq 0}$ such that

(a) $d(p, q) = d(q, p)$ for all $p, q \in X$,

(b) $d(p, q) = 0$ if and only if $p = q$, and

(c) $d(p, q) \leq d(p, r) + d(r, q)$ for all $p, q, r \in X$.

The last property is called the *triangle inequality*:



Now we can translate our definition of continuity and convergence in this abstract setting:

Definition 1.3. Let X and Y be sets equipped with metrics d_X and d_Y.

(a) A map $f\colon X \to Y$ is *continuous* at a point $p \in X$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that for all $q \in X$ with $d_X(p, q) < \delta$ we have $d_Y(f(p), f(q)) < \epsilon$.

(b) A sequence of points $p_1, p_2, p_3, \ldots \in X$ *converges* to a limit $\ell \in X$ if for every $\epsilon > 0$ there is a natural number N such that for all $n \geq N$ we have $d_X(p_n, \ell) < \epsilon$.

We will often say "let (X, d) be a metric space," which means that X is a set and d is a metric on it. A set will typically admit many metrics, or to put it another way, we can have many metric spaces with the same underlying set. Here are some of the examples that we will be interested in.

Example 1.4. *Three metrics on* $\mathbb{R}^n$.

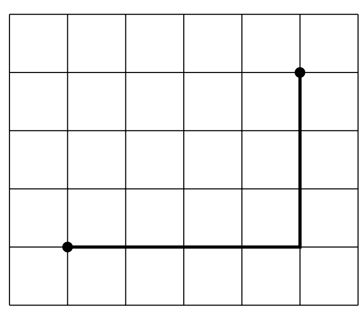
(a) The Euclidean metric: given two points $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, we define

$$d_2(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}.$$

(b) The taxicab metric

$$d_1(\mathbf{x}, \mathbf{y}) = |x_1 - y_1| + |x_2 - y_2| + \cdots + |x_n - y_n|.$$

The name alludes to the driving distance between two points in a city whose roads are laid out on a grid, like Manhattan or Eugene.



(c) The square metric on $\mathbb{R}^n$:

$$d_\infty(\mathbf{x}, \mathbf{y}) = \max\{|x_1 - y_1|, \, |x_2 - y_2|, \, \ldots, \, |x_n - y_n|\}.$$

The reason for the name will become clear in Exercise 1.1.

On $\mathbb{R}^1$ these three metrics are all the same, but when $n \geq 2$ they are all different: for example, they give different distances between $(0, 0, \ldots, 0)$ and $(1, 1, \ldots 1)$. We will eventually see, however, that they all have the same continuous maps and convergent sequences.

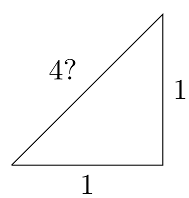
Example 1.5. *Interpolating between those three metrics, and a non-example.*
For $p > 0$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ we can define

$$d_p(\mathbf{x}, \mathbf{y}) = \big(|x_1 - y_1|^p + |x_2 - y_2|^p + \cdots + |x_n - y_n|^p\big)^{1/p}.$$

We see that $p = 1$ gives the taxicab metric, $p = 2$ gives the Euclidean metric, and it is interesting to think about how letting $p \to \infty$ gives the square metric. We see that d_p satisfies the first two properties of Definition 1.2, and it turns out to satisfy the last property (the triangle inequality) if $p \geq 1$, although this is not obvious. On the other hand, if $0 < p < 1$ then the triangle inequality fails: for example, with $p = 1/2$ we have

$$\begin{aligned}
d_{1/2}((0,0),(0,1)) &= \big(|0-0|^{1/2} + |0-1|^{1/2}\big)^2 = (0+1)^2 = 1, \\
d_{1/2}((0,1),(1,1)) &= \big(|0-1|^{1/2} + |1-1|^{1/2}\big)^2 = (0+1)^2 = 1, \\
d_{1/2}((0,0),(1,1)) &= \big(|0-1|^{1/2} + |0-1|^{1/2}\big)^2 = (1+1)^2 = 4,
\end{aligned}$$

but 4 is bigger than $1 + 1$.



Example 1.6. To illustrate how permissive the definition of a metric is, we introduce the *SNCF metric* on $\mathbb{R}^2$:

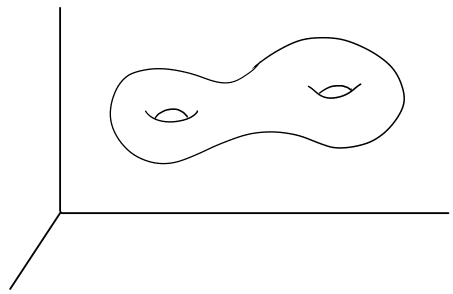
$$d_*(p,q) = \begin{cases} d_2(p,q) & \text{if } p \text{ and } q \text{ lie on the same line through the origin, or} \\ d_2(p,0) + d_2(0,q) & \text{otherwise,} \end{cases}$$

where d_2 is the Euclidean metric. You should convince yourself that this satisfies the triangle inequality.

SNCF stands for *Société nationale des chemins de fer français*, the French national railway company; the joke is that if you want to go from, say, Dijon to Bordeaux, you might as well take the train up to Paris and back down. In the UK, it is called the British Rail metric; in Eugene, the LTD bus metric.

Example 1.7. *The induced metric on a subset.*

If Y is a subset of X, then a metric $d\colon X \times X \to \mathbb{R}$ induces a metric on Y, just by restricting the domain of d to $Y \times Y \subset X \times X$: that is, by declaring that the distance between two points in the subspace Y is the same as it was in the ambient space X. So for example if $X = \mathbb{R}^3$ and Y is a surface,



then we can get metrics on Y by restricting the Euclidean metric, the taxicab metric, or the square metric from $\mathbb{R}^3$.

This surface carries other interesting metrics as well: for example, we could declare the distance between two points on the surface to be the length of the shortest path between them on the surface, which will be different from the shortest path through the ambient space. That metric belongs to a subject called Riemannian geometry, and is beyond the scope of this course.

Example 1.8. *Some spaces of functions.*

Let $C([0,1])$ be the set whose elements are continuous, real-valued functions $f\colon [0,1] \to \mathbb{R}$. This set is much too big to visualize like we do $\mathbb{R}^2$ or $\mathbb{R}^3$, but we want to regard its elements as points in a huge space and consider different ways of measuring the distance between them. There are many interesting metrics on $C([0,1])$, for example the L^1 metric

$$d_1(f,g) = \int_0^1 |f(x) - g(x)|\, dx,$$

the L^2 metric

$$d_2(f,g) = \sqrt{\int_0^1 |f(x) - g(x)|^2\, dx},$$

the L^∞ or sup metric

$$d_\infty(f,g) = \sup_{x \in [0,1]} |f(x) - g(x)|,$$

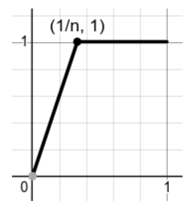
and the L^p metrics for $p \geq 1$, which interpolate between these:

$$d_p(f, g) = \left(\int_0^1 |f(x) - g(x)|^p \, dx \right)^{1/p}.$$

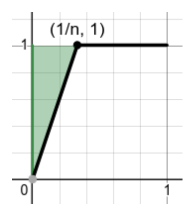
It is interesting to think about how these metrics are analogous to the metrics on $\mathbb{R}^n$ that we named d_1, d_2, d_∞, and d_p a moment ago. But whereas those metrics on $\mathbb{R}^n$ will all turn out to give the same continuous maps and convergent sequences, these ones do not:

(a) *A sequence in $C([0, 1])$ that converges in the L^1 metric but not in the sup metric.*

For $n = 1, 2, 3, \ldots$, consider the function $f_n : [0, 1] \to \mathbb{R}$ that goes piecewise linearly from $f(0) = 0$ to $f(1/n) = 1$ to $f(1) = 1$:



I claim that the sequence $f_1, f_2, f_3, \ldots$ converges to the constant function $g(x) = 1$ in the L^1 metric, but not in the sup metric. On the one hand, $d_1(f_n, g) = \int_0^1 |f_n - g|$ is the area of the triangle shown,



so that's $1/2n$, which goes to zero as $n \to \infty$, so $f_n \to g$ in the L^1 metric. On the other hand, $d_\infty(f_n, g) = \sup_{x \in [0,1]} |f_n(x) - g(x)|$ is equal to 1 for all n, which does not go to zero as $n \to \infty$, so $f_n \not\to g$ in the sup metric.

You can check that the same thing happens with $f_n(x) = x^n$ and $g(x) = 0$.

You can unpack the definitions to see that a sequence of functions converges in the sup metric if and only if it converges *uniformly*.

(b) *Integration gives a map from $C([0,1])$ to $\mathbb{R}$ that is continuous in both the sup metric and the L^1 metric.*

Consider the map $\Phi\colon C([0,1]) \to \mathbb{R}$ given by $\Phi(f) = \int_0^1 f(x)\,dx$.

First I claim that Φ is continuous if $C([0,1])$ is given the sup metric (and the target space $\mathbb{R}$ is given the usual Euclidean metric). Unpacking the definitions, this means that for every $f \in C([0,1])$ and every $\epsilon > 0$, there is a $\delta > 0$ such that for all $g \in C([0,1])$ with $d_\infty(f,g) < \delta$ we have $|\Phi(f) - \Phi(g)| < \epsilon$. So let $f \in C([0,1])$ and $\epsilon > 0$ be given, and choose some $\delta < \epsilon$, say $\delta = \epsilon/2$. If $d_\infty(f,g) < \delta$, then for all $x \in [0,1]$ we have $|f(x) - g(x)| < \delta$, so

$$|\Phi(f) - \Phi(g)| = \left| \int_0^1 f(x)\,dx - \int_0^1 g(x)\,dx \right|$$
$$= \left| \int_0^1 (f(x) - g(x))\,dx \right| \leq \int_0^1 |f(x) - g(x)|\,dx \leq \int_0^1 \delta\,dx = \delta < \epsilon,$$

which is what we wanted.

Next I claim that Φ is also continuous in the L^1 metric. Let $f \in C([0,1])$ and $\epsilon > 0$ be given, and take $\delta = \epsilon$. If $d_1(f,g) < \delta$, then

$$|\Phi(f) - \Phi(g)| = \left| \int_0^1 f(x)\,dx - \int_0^1 g(x)\,dx \right|$$
$$= \left| \int_0^1 (f(x) - g(x))\,dx \right| \leq \int_0^1 |f(x) - g(x)|\,dx = d_1(f,g) < \delta = \epsilon,$$

which is what we wanted.

(c) *Evaluation at $x = 0$ gives a map from $C([0,1])$ to $\mathbb{R}$ that is continuous in the sup metric but not in the L^1 metric.*

Consider the map $\Psi\colon C([0,1]) \to \mathbb{R}$ given by $\Psi(f) = f(0)$.

On the one hand, I claim that Ψ is continuous in the sup metric. Let $f \in C([0,1])$ and $\epsilon > 0$ be given, and take $\delta = \epsilon$. If $d_\infty(f,g) < \delta$, then for all $x \in [0,1]$ we have $|f(x) - g(x)| < \delta$, so

$$|\Psi(f) - \Psi(g)| = |f(0) - g(0)| < \delta = \epsilon,$$

as desired.

On the other hand, I claim that Ψ is *not* continuous in the L^1 metric. Take the sequence $f_1, f_2, f_3, \dots$ from part (a) above, which converges in the L^1 metric to the constant function g. If Ψ were continuous in the L^1 metric, then the sequence $\Psi(f_1), \Psi(f_2), \Psi(f_3), \dots$ in $\mathbb{R}$ would converge to $\Psi(g)$ by Exercise 1.11 below, but in fact we have $\Psi(f_n) = 0$ for all n while $\Psi(g) = 1$.

Our definition of the sup metric implicitly uses the fact that a continuous function on $[0,1]$ is bounded: otherwise $\sup|f-g|$ might be infinite. You have probably seen this proved in a first course in real analysis; we will assume it for now, and prove it properly when we come to compactness.

Example 1.9. There are many other function spaces. For example, we can take the set $C^1([0,1])$ of functions $f\colon [0,1] \to \mathbb{R}$ that are continuously differentiable, meaning that the derivative f' exists and is continuous, and give it the metric

$$d(f,g) = \sup|f-g| + \sup|f'-g'|,$$

called the C^1 *metric*. Or we can give it one of the *Sobolev metrics*

$$d(f,g) = \Big((d_p(f,g))^p + (d_p(f',g'))^p\Big)^{1/p},$$

where d_p is the L^p metric from Example 1.8 and $p \geq 1$. More generally, we can take the set $C^k([0,1])$ of functions whose first k derivatives exist and are continuous, and give it a similarly-defined C^k metric, Sobolev metrics, and many others. We can generalize further to functions of several variables, and on and on. These function spaces are useful in studying solutions to partial differential equations.

We conclude this section with a key fact relating continuous maps and convergent sequences, leaving part of the proof as an exercise.

Proposition 1.10. *Let (X, d_X) and (Y, d_Y) be metric spaces. Then a map $f\colon X \to Y$ is continuous at a point $\ell \in X$ if and only if for every sequence $p_1, p_2, p_3, \dots$ in X that converges to ℓ, the sequence $f(p_1), f(p_2), f(p_3), \dots$ in Y converges to $f(\ell)$.*

Proof. One direction – if f is continuous at ℓ then it takes any sequence converging to ℓ to a sequence converging to $f(\ell)$ – is Exercise 1.11 below. For the other direction, let us argue that if f is *not* continuous at ℓ, then there is a sequence $p_1, p_2, \ldots \in X$ that converges to ℓ, but $f(p_1), f(p_2), \ldots \in Y$ does not converge to $f(\ell)$.

To say that f is not continous at ℓ means that there is an $\epsilon > 0$ such that for every $\delta > 0$, there is a point $p \in X$ with $d_X(p, \ell) < \delta$ but $d_Y(f(p), f(\ell)) \geq \epsilon$. So for each $n = 1, 2, 3, \ldots$, we can take $\delta = 1/n$ and get a point p_n such that $d_X(p_n, \ell) < 1/n$ but $d(f(p_n), f(\ell)) \geq \epsilon$. Then the sequence $p_1, p_2, p_3, \ldots$ converges to ℓ, but $f(p_1), f(p_2), f(p_3), \ldots$ does not converge to $f(\ell)$. (You may want to write out the details of these last two claims as well.) $\square$

Exercises.

1.1. (a) For each of the three metrics in Example 1.4, sketch the open ball of some radius $r > 0$ around the origin in $\mathbb{R}^2$:

$$B_r(0) = \{(x, y) \in \mathbb{R}^2 : d((x, y),\, 0) < r\}.$$

(b) For one of the three metrics (your choice), prove or give a counterexample to the following statement: a sequence of points (x_1, y_1), (x_2, y_2), (x_3, y_3), $\ldots \in \mathbb{R}^2$ converges to a limit (x, y) if and only if $x_n \to x$ and $y_n \to y$ separately, as sequences in $\mathbb{R}$ with the usual metric.

(c) Why is

$$d(\mathbf{x}, \mathbf{y}) = \min\{|x_1 - y_1|, |x_2 - y_2|, \ldots, |x_n - y_n|\}$$

not a metric on $\mathbb{R}^n$? [Part (c) was not part of Homework 1, I added it to the notes later.]

1.2. The SNCF metric on $\mathbb{R}^2$ was introduced in Example 1.6.

(a) Give an example of a sequence that converges in the Euclidean metric but not in the SNCF metric.

(b) Prove that every sequence that converges in the SNCF metric converges in the Euclidean metric.

1.3. Consider the following silly metric on $\mathbb{R}^2$:

$$d((x_1, y_1),\ (x_2, y_2)) = \begin{cases} |y_1 - y_2| & \text{if } x_1 = x_2 \\ |y_1 - y_2| + 1 & \text{if } x_1 \neq x_2. \end{cases}$$

(a) Prove that d is a metric, that is, it has the three properties listed in Definition 1.2.

(b) Sketch the open balls of radius $1/2$, 1, and 2 around the origin in this metric.

(c) Give an example of a sequence that converges in the Euclidean metric d_2 but not in our silly metric d.

(d) Prove that every sequence that converges in d also converges d_2.

1.4. In Example 1.8(a) we saw a sequence in $C([0, 1])$ that converges in the L^1 metric but not in the sup metric. Prove that the reverse cannot happen: every sequence that converges in the sup metric converges in the L^1 metric.

1.5. Let (X, d) be a metric space. Prove the *reverse triangle inequality*:

$$|d(p, q) - d(p, r)| \leq d(q, r)$$

for all $p, q, r \in X$. Include an appropriate picture.

1.6. Let (X, d_X) and (Y, d_Y) and (Z, d_Z) be metric spaces. Let $f\colon X \to Y$ be continuous at a point $p \in X$, and let $g\colon Y \to Z$ be continuous at $f(p)$. Prove that $g \circ f$ is continuous at p.

1.7. Let (X, d) be a metric space, and fix a point $a \in X$. Prove that the map $f\colon X \to \mathbb{R}$ given by $f(p) = d(p, a)$ is continuous.

Hint: Use the triangle inequality.

1.8. Let X be any set, and let d_X be the *discrete metric*

$$d_X(p, q) = \begin{cases} 0 & \text{if } p = q, \text{ or} \\ 1 & \text{if } p \neq q. \end{cases}$$

(a) Prove that d_X is a metric.

(b) Let (Y, d_Y) be another metric space (not necessarily discrete). Prove that every map $f\colon X \to Y$ is continuous.

(c) Prove that a sequence $p_1, p_2, p_3, \ldots \in X$ converges in the discrete metric if and only if it is eventually constant.

1.9. Let $p_1, p_2, p_3, \ldots$ be a sequence in a metric space (X, d), and suppose that $p_n \to \ell$ and $p_n \to \ell'$ for two points $\ell, \ell' \in X$. Prove that $\ell = \ell'$.

Hint: Prove that $d(\ell, \ell') < \epsilon$ for every $\epsilon > 0$, using the triangle inequality. Then argue that this implies $\ell = \ell'$.

1.10. Let $p_1, p_2, p_3, \ldots$ and $p'_1, p'_2, p'_3, \ldots$ be two sequences in a metric space (X, d). Prove that if $p_n \to \ell$ and $d(p_n, p'_n) \to 0$ as $n \to \infty$, then $p'_n \to \ell$.

1.11. (One direction of Proposition 1.10.) Let (X, d_X) and (Y, d_Y) be metric spaces, let $p_1, p_2, p_3, \ldots$ be a sequence that converges to a point ℓ in X, and let $f\colon X \to Y$ be continuous at ℓ. Prove that the sequence $f(p_1), f(p_2), f(p_3), \ldots$ converges to $f(\ell)$ in Y.

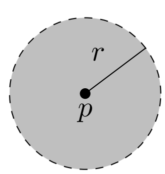
1.12. Let
$$W = \{1, \tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{4}, \ldots, 0\}$$
with the metric induced from the usual one on $\mathbb{R}$. Let (X, d_X) be another metric space. Given a sequence $p_1, p_2, p_3 \ldots \in X$ and a point $\ell \in X$, prove that the map $f\colon W \to X$ defined by
$$\begin{cases} f(\frac{1}{n}) = p_n, \\ f(0) = \ell \end{cases}$$
is continuous if and only if $p_n \to \ell$.

2 Open and Closed Sets

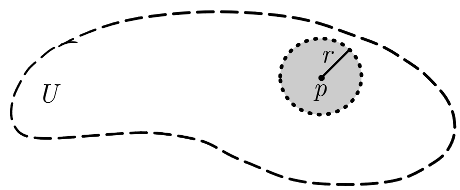
Let us fix a metric space (X, d) for the whole section.

Definition 2.1. The *open ball* of radius r around a point $p \in X$ is
$$B_r(p) = \{q \in X : d(p, q) < r\} \subset X.$$

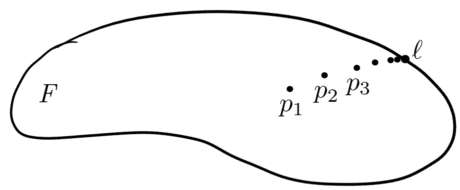
In $\mathbb{R}^2$ with the Euclidean metric, this looks like



Definition 2.2. A subset $U \subset X$ is *open* if for every $p \in U$ there is an $r > 0$ such that $B_r(p) \subset U$.

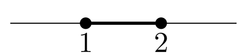


Definition 2.3. A subset $F \subset X$ is *closed* if for every sequence of points $p_1, p_2, p_3, \ldots \in F$ converging to a limit $\ell \in X$, we have $\ell \in F$.

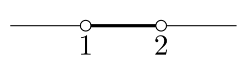


Thus an open subset is one where you can move around a little bit without leaving the subset, and a closed subset is one where you can't get out by taking the limit of a sequence. The letter F is from the French *fermé*, closed. The letter U seems to be from the German *Umgebung*, neighborhood, as in the early texts by Hausdorff [2, VII §1] and Tietze [3].

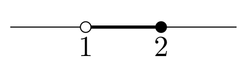
Example 2.4. In $\mathbb{R}$ in the usual topology, the interval $[1, 2]$ is closed, but not open – any open ball around 1 or 2 spills out.



The interval $(1, 2)$ is open but not closed – the sequence $\frac{3}{2}, \frac{4}{3}, \frac{5}{4}, \ldots$ converges to 1 which is not in the set.



The interval $(1, 2]$ is neither open nor closed.

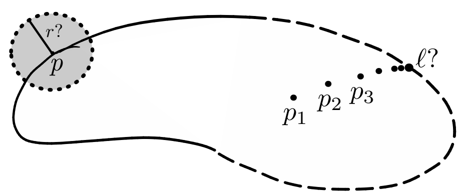


The French write $]1, 2[$ rather than $(1, 2)$, reserving the latter for an ordered pair or point in the plane; similarly they write $]1, 2]$ rather than $(1, 2]$.

Example 2.5. The rational numbers $\mathbb{Q} \subset \mathbb{R}$ are neither open nor closed in the usual metric. On the one hand, if x is rational and $r > 0$ then the ball $B_r(x)$, which is just the interval $(x - r, x + r)$, contains irrational numbers. On the other hand, a sequence of rational numbers can converge to an irrational number: for example

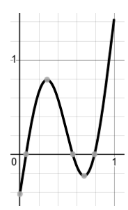
$$1,\ 1.4,\ 1.41,\ 1.414,\ 1.4142,\ 1.41421,\ 1.414213 \to \sqrt{2}$$

Example 2.6. The picture in Definition 2.2 shows a subset of $\mathbb{R}^2$ that is open but not closed in the Euclidean metric, and the one in Definition 2.3 shows one that is closed but not open. Here is one that is neither open nor closed:



You should get the feeling that "most" subsets are neither open or closed.

Example 2.7. Let $X = C^1([0,1])$, the set of continuously differentiable functions, with the C^1 metric from Example 1.9. The subset consisting of functions with simple roots – that is, those for which $f'(x) \neq 0$ whenver $f(x) = 0$ – turns out to be open. Exercise 2.3 asks you to prove this, but it agrees with our intuition that if we take a function with simple roots and wiggle it a bit, then it still has simple roots:



This is a first taste of what is called *transversality*.

But the set of functions with simple roots is not closed: for example, the functions

$$f_n(x) = \left(x - \tfrac{1}{2} - \tfrac{1}{n}\right)\left(x - \tfrac{1}{2} + \tfrac{1}{n}\right)$$

all have simple roots (at $\frac{1}{2} \pm \frac{1}{n}$), but they converge to $g(x) = (x - \frac{1}{2})^2$ which does not.

Example 2.8. Let $X = C([0,1])$ with the sup metric, and fix a subset $A \subset [0,1]$. The subset consisting of functions that vanish on A – that is, those with $f(x) = 0$ for all $x \in A$ – turns out to be closed in the sup metric. Exercise 2.4 asks you to prove this. But it is not closed in the L^1 metric, as we can see from Example 1.8(a).

Example 2.9. In a discrete metric (Exercise 1.8), every subset is both open and closed.

Here are some basic properties of closed sets. Exercise 2.6 asks you to prove the analogous properties of open sets.

Proposition 2.10. *If $F, G \subset X$ are closed, then the union $F \cup G$ is again closed.*

Proof. Let $p_1, p_2, p_3, \ldots \subset F \cup G$ be a sequence converging to a limit $\ell \in X$. By the pigeonhole principle, either there are infinitely many n such that $p_n \in F$, or infinitely many such that $p_n \in G$, or both. In the first case, the subsequence consisting of p_ns that lie in F still converges to ℓ; because F is closed, we have $\ell \in F$, so $\ell \in F \cup G$. Similarly, in the second case we get $\ell \in F \cup G$. $\square$

By induction, if we have finitely many closed sets $F_1, F_2, \ldots, F_n \subset X$, then their union $F_1 \cup F_2 \cup \cdots \cup F_n$ is again closed. But we could have a countable collection of closed sets whose union is not countable: for example, if for $n = 1, 2, 3, \ldots$ we set $F_n = [\frac{1}{n}, 1] \subset \mathbb{R}$, then each F_n is closed in the usual metric, but

$$F_1 \cup F_2 \cup F_3 \cup \cdots = (0,1]$$

is not closed.

On the other hand, an arbitrary (even uncountable) intersection of closed sets is closed:

Proposition 2.11. *Let I be a set, and suppose that for each $i \in I$ we have a closed set $F_i \subset X$. Then the intersection $\bigcap_{i \in I} F_i$ is again closed.*

Proof. Let $p_1, p_2, p_3, \ldots \in \bigcap F_i$ be a sequence converging to a limit $\ell \in X$. For every $i \in I$ we have $p_1, p_2, p_3, \ldots \in F_i$. Because F_i is closed, we have $\ell \in F_i$. Because this is true for every $i \in I$, we have $\ell \in \bigcap F_i$. $\square$

Open and closed sets are dual to one another, in the following sense:

Proposition 2.12. *A subset $A \subset X$ is closed if and only if the complement $X \setminus A$ is open.*

Proof. We will prove that A is *not* closed if and only if the complement $X \setminus A$ is *not* open.

First suppose that A is not closed, meaning that there is a sequence $p_1, p_2, p_3, \ldots \in A$ that converges to a limit $\ell \in X \setminus A$. For every $\epsilon > 0$ there is an N such that for all $n \geq N$, we have $d(p_n, \ell) < \epsilon$. Thus the ball $B_\epsilon(\ell)$ contains points $p_N, p_{N+1}, \ldots$ that are in A, so the ball is not contained in $X \setminus A$. Thus $X \setminus A$ is not open, because there is no ball around ℓ that is contained in $X \setminus A$.

Conversely, suppose that $X \setminus A$ is not open, meaning that there is a point $p \in X \setminus A$ such that for every $r > 0$, the ball $B_r(p)$ is not contained in $X \setminus A$, that is, it meets A. For each $n = 1, 2, 3, \ldots$, take $r = 1/n$, and choose a point $q_n \in B_{1/n}(p) \cap A$. Then the sequence $q_1, q_2, q_3, \ldots$ is contained in A, but $d(q_n, p) < 1/n$, so $q_n \to p$ which is not in A. Thus A is not closed. □

Exercises.

2.1. Let $X = \mathbb{R}^2$ with the Euclidean metric. Sketch the subset

$$A = \{(x, y) \in \mathbb{R}^2 : x \neq 0 \text{ or } y = 0\}.$$

Prove that A is neither open nor closed.

2.2. Let $X = \mathbb{Q}$ with the metric induced from the usual one on $\mathbb{R}$: that is, $d(x, y) = |x - y|$ for all $x, y \in \mathbb{Q}$, but we're thinking about $\mathbb{Q}$ in itself and forgetting about the rest of $\mathbb{R}$.

(a) Prove that the subset

$$\{x \in \mathbb{Q} : x^2 < 1\}$$

is open but not closed.

(b) Prove that the subset

$$\{x \in \mathbb{Q} : x^2 < 2\}$$

is both open and closed.
(You may use the fact that $\sqrt{2}$ is irrational without proving it.)

2.3. Let $U \subset C^1([0,1])$ be the set of functions with simple roots as in Example 2.7. Prove that U is open in the C^1 metric.

Hint: For a given $f \in U$, take the ball of radius

$$r = \inf_{x\in[0,1]} \big(|f(x)| + |f'(x)|\big) .$$

2.4. Let $A \subset [0,1]$, and let $F \subset C([0,1])$ be the set of continuous functions that vanish on A as in Example 2.8. Prove that F is closed in the sup metric.

Hint: One possibility is to use Proposition 1.10 together with Example 1.8(c), which is stated for evaluation at $x = 0$ but which we can see is equally valid for evaluation at any $x \in [0,1]$.

2.5. (a) For a point $p \in X$ and a radius $r > 0$, the open ball $B_r(p) \subset X$ was defined in Definition 2.1. Prove that it is open.

(b) Define the closed ball

$$\bar{B}_r(p) = \{q \in X : d(p,q) \leq r\}.$$

Prove that it is closed.

Hint: You can prove it directly, or you can use Proposition 2.12.

2.6. Without using Proposition 2.12,

(a) Prove that if $U, V \subset X$ are open, then the intersection $U \cap V$ is again open.

(b) Give an example of countably many open sets $U_1, U_2, U_3, \ldots \subset X$ such that their intersection $U_1 \cap U_2 \cap U_3 \cap \cdots$ is not open.

(c) Let I be a set, and suppose that for each $i \in I$ we have an open set $U_i \subset X$. Prove that the union $\bigcup_{i\in I} U_i$ is again open.
(Don't assume that the index set I is countable!)

2.7. Given a subset $A \subset X$, a point $p \in X$ is called a *limit point* of A if for every $r > 0$ there is a point $q \in A \cap B_r(p)$ with $q \neq p$.

Prove that A is closed if and only if it contains all its limit points.

(Some authors take this as the definition of a closed set.)

2.8. In Example 1.4 we saw three different metrics on $\mathbb{R}^2$. Prove one of the following:

(a) A subset $A \subset \mathbb{R}^2$ is open in the Euclidean metric if and only if it is open in the taxicab metric.

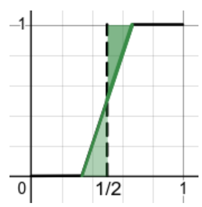
(b) A subset $A \subset \mathbb{R}^2$ is open in the Euclidean metric if and only if it is open in the square metric.

(c) A subset $A \subset \mathbb{R}^2$ is open in the taxicab metric if and only if it is open in the square metric.

3 Completeness

We have called a subset $F \subset X$ *closed* if it contains all the limits it should – or at least, all the limits that X knows about. Sometimes, however, the metric space X itself has sequences that seem like they ought to converge, but the limit is missing. For example, take $X = C([0,1])$ with the L^1 metric, and for $n = 2, 3, 4, \ldots$, consider the function $f_n \colon [0,1] \to \mathbb{R}$ that goes piecewise linearly from $f(0) = 0$ to $f(\frac{1}{2} - \frac{1}{n}) = 0$ to $f(\frac{1}{2} + \frac{1}{n}) = 1$ to $f(1) = 1$:

As $n \to \infty$, we see that $f_n(x)$ looks more and more the step function $g(x)$ that jumps from 0 to 1 at $x = 1/2$. The sequence does not converge to g in the sup metric, but it would in the L^1 metric, if g were in $C([0,1])$: the distance between f_n and g would be the area of the two triangles shown,



which is $\frac{1}{2n} \to 0$. But g is not in $C([0,1])$. We will name this problem by saying that the L^1 metric on $C([0,1])$ is *incomplete*.

For a more familiar example, the sequence

$$1,\ 1.4,\ 1.41,\ 1.414,\ 1.4142,\ 1.41421,\ 1.414213,\ \ldots \in \mathbb{Q}$$

wants to approach a limit, namely $\sqrt{2}$, but the limit is missing, so we have to complete the rational numbers $\mathbb{Q}$ to get the real numbers $\mathbb{R}$.

To formalize this, we introduce the definition of a Cauchy sequence,* which tries to say that the sequence converges without referring to the limit, because the limit might be missing from our metric space.

Definition 3.1. A sequence $p_1, p_2, p_3, \ldots$ in a metric space (X, d) is *Cauchy* if for every $\epsilon > 0$ there is a natural number N such that for all $m, n \geq N$ we have $d(p_m, p_n) < \epsilon$.

In contrast to Definition 1.3(b) of a convergent sequence, where the tails of the sequence get arbitrarily close to a limit point ℓ, here the tails just get close to themselves.

Proposition 3.2. *Let (X, d) be a metric space. If a sequence $p_1, p_2, p_3, \ldots$ converges to a limit ℓ, then it is Cauchy.*

Proof. Let $\epsilon > 0$ be given. Because the sequence converges to ℓ, there is a natural number N such that $n \geq N$ implies $d(p_n, \ell) < \epsilon/2$. Thus if $m, n \geq N$ then the triangle inequality gives

$$d(p_m, p_n) \leq d(p_m, \ell) + d(\ell, p_n) < \epsilon/2 + \epsilon/2 = \epsilon. \qquad \square$$

Definition 3.3. A metric space is *complete* if every Cauchy sequence converges.

Proposition 3.4. $\mathbb{R}$ *is complete in the usual metric.*

You have probably seen this proved in a first course in real analysis, but we review the proof, which relies on the fact that a bounded set in $\mathbb{R}$ has a supremum (least upper bound) and an infimum (greatest lower bound). That is, the completeness of $\mathbb{R}$ as a metric space follows from its completeness as an ordered set.

---

*Named for Augustin-Louis Cauchy, 1789–1857, and pronounced ko-shee, not kaw-shee. It's difficult for English speakers to avoid stressing one syllable or the other, so American speakers tend say ko-SHEE, while British speakers tend to say KO-shee. The same thing happens with ballet, garage, and many other words borrowed from French.

Proof of Proposition 3.4. Let $x_1, x_2, x_3, \ldots \in \mathbb{R}$ be a Cauchy sequence.

Exercise 3.3 asks you to prove that a Cauchy sequence is bounded: thus there is an $M \in \mathbb{R}$ such that $|x_n| \le M$ for all n.

Take the limit superior and limit inferior

$$\ell = \liminf_{n\to\infty} x_n \qquad \text{and} \qquad L = \limsup_{n\to\infty} x_n,$$

which are defined as follows. For $n = 1, 2, 3, \ldots$, let

$$a_n = \inf\{x_n, x_{n+1}, x_{n+2} \ldots\} \quad \text{and} \quad A_n = \sup\{x_n, x_{n+1}, x_{n+2}, \ldots\},$$

which exist because our sequence is bounded. We see that

$$-M \le a_n \le x_n \le A_n \le M. \tag{3.1}$$

We define

$$\ell = \sup\{a_1, a_2, a_3 \ldots\} \qquad \text{and} \qquad L = \sup\{A_1, A_2, A_3, \ldots\}.$$

The sequence $a_1, a_2, a_3, \ldots$ is non-decreasing, because each term is the infimum of a smaller set than the one before, and a bounded non-decreasing sequence converges to its supremum, so $a_n \to \ell$ as $n \to \infty$. Similarly we have $A_n \to L$ as $n \to \infty$.

We want to prove that the sequence $x_1, x_2, x_3, \ldots \cdot$ converges. By inequalities (3.1) and the squeeze theorem, it is enough to prove that $\ell = L$, or equivalently, that $|A_n - a_n| \to 0$ as $n \to \infty$. Let $\epsilon > 0$ be given. Because $x_1, x_2, \ldots$ is a Cauchy sequence, there is an N such that for all $m, n \ge N$ we have $|x_m - x_n| < \epsilon/3$. Because a_N was defined as an infimum, there is some $m \ge N$ such that $x_m < a_N + \epsilon/3$; otherwise $a_N + \epsilon/3$ would have been a greater lower bound for the set $\{x_N, x_{N+1}, \ldots\}$. Similary, there is an $n \ge N$ such that $x_n > A_N - \epsilon/3$. Now by the triangle inequality we have

$$|A_N - a_N| \le |A_N - x_n| + |x_n - x_m| + |x_m - a_N| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.$$

Because the As are non-increasing and the as are non-decreasing, for all $\nu \ge N$ we have $|A_\nu - a_\nu| \le |A_N - a_N| < \epsilon$, which is what we wanted. $\square$

Example 3.5. On the other hand, $\mathbb{R}$ is *not* complete in the metric

$$d(x, y) = |\arctan x - \arctan y|,$$

because the sequence $1, 2, 3, \dots$ is Cauchy but does not converge. To see that it is Cauchy, note that the sequence $\arctan 1, \arctan 2, \arctan 3, \dots$ converges to $\pi/2$ in the usual metric, so it is Cauchy in the usual metric, so for every $\epsilon > 0$ there is an N such that if $m, n \geq N$ then $|\arctan m - \arctan n| < \epsilon$. To see that it does not converge, note that the distance from $1, 2, 3, \dots$ to any $\ell \in \mathbb{R}$ approaches $\pi/2 - \arctan(\ell)$, and in particular does not go to zero.

If we wanted to make $\mathbb{R}$ complete in this metric, we would need to add two more points at $\pm\infty$. We will return to the idea of completion soon.

One can show that $\mathbb{R}$ with this metric has the same continuous functions, convergent sequences, and open and closed sets as it does in the usual metric. Later we will see that continuity and convergence can be defined purely in terms of open sets, but this example shows that completeness cannot.

Example 3.6. $\mathbb{R}^n$ is complete in any of the metrics from Example 1.4. To see this, let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots$ be a sequence in $\mathbb{R}^n$ that is Cauchy under one of the three metrics. Exercise 1.1(b) asked you to prove that the sequence of points in $\mathbb{R}^2$ converges in one of the three metrics if and only if each of its coordinates converges in the usual metric, and the same proof applies to $\mathbb{R}^n$. So if we let the i^{th} coordinate of $\mathbf{x}_n$ be called $x_{n,i}$, then it is enough to prove that the sequence $x_{1,i}, x_{2,i}, x_{3,i}, \dots$ is Cauchy in $\mathbb{R}$ in the usual metric. This follows from the fact that

$$|x_{m,i} - x_{n,i}| \leq d(\mathbf{x}_m, \mathbf{x}_n),$$

in any of the three metrics.

Example 3.7. *Poincaré's hyperbolic ball and half-space.*

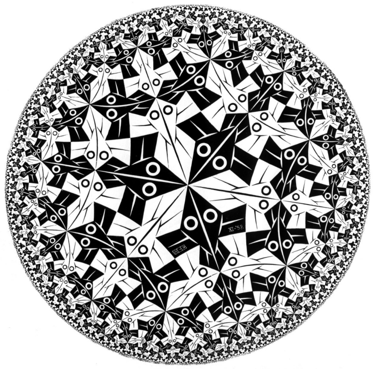
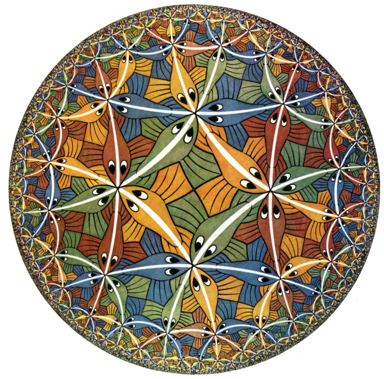
Consider the open unit ball

$$B = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}| < 1\}.$$

The Euclidean metric that B inherits as a subset of $\mathbb{R}^n$ is not complete: a sequence that approaches a point on the boundary sphere will be Cauchy, but will not converge. On the other hand, the *hyperbolic metric*

$$d(\mathbf{x}, \mathbf{y}) = \cosh^{-1}\left(1 + 2\frac{|\mathbf{x} - \mathbf{y}|^2}{(1 - |\mathbf{x}|^2)(1 - |\mathbf{y}|^2)}\right)$$

turns out to be complete. Here are two images from a series of four by M. C. Escher set in the hyperbolic disc; any two fish have the same length in the hyperbolic metric.

As in Example 3.5, the hyperbolic and Euclidean metrics have the same continuous functions, convergent sequences, and open and closed sets, but one is complete while the other is not.

Similarly, the upper half-space

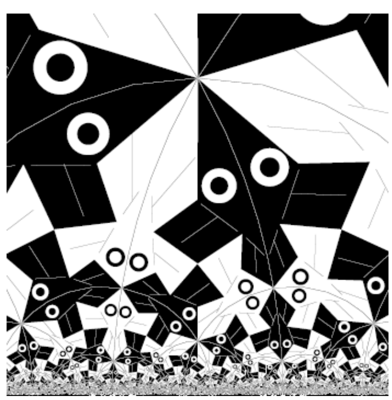
$$H = \{\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n : x_n > 0\}$$

is not complete in the Euclidean metric, but is complete in the hyperbolic metric

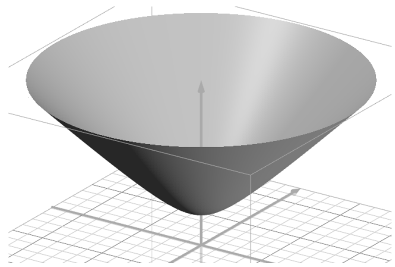
$$d(\mathbf{x}, \mathbf{y}) = 2 \sinh^{-1} \frac{|\mathbf{x} - \mathbf{y}|}{2\sqrt{x_n y_n}}.$$

It is interesting to check that the sequence $(0, 0, \dots, 0, \frac{1}{n})$, which is Cauchy in the Euclidean metric, is not Cauchy in this metric.

There is a bijection between the ball and the half-space that turns one hyperbolic metric into the other. Here is the image of Escher's first print under that isometry, reproduced from [1]:



Moreover there are bijections from either the ball or the half-space to the hyperboloid $x_{n+1} = \sqrt{1 + x_1^2 + \ldots + x_n^2}$ in $\mathbb{R}^{n+1}$,



under which the hyperbolic metrics above correspond to the Riemannian distance function given by the shortest path along the hyperboloid, as mentioned briefly in Example 1.7.

These hyperbolic spaces are important in differential geometry, complex analysis, and number theory, and historically in demonstrating the logical independence of Euclid's parallel postulate from his other four.

Example 3.8. Given a prime number p, the *p-adic metric* on $\mathbb{Q}$ is defined as follows. For $x, y \in \mathbb{Q}$ with $x \neq y$, we can write

$$|x - y| = p^v \cdot \frac{r}{s},$$

where $v, r, s \in \mathbb{Z}$ and r and s are not divisible by p. Then we define $d_p(x, y) = p^{-v}$, or if $x = y$ then $d_p(x, y) = 0$.* Exercise 3.2 asks you to check that this is a metric and explore some numerical examples.

Convergent sequences in the p-adic metric can look bizarre: for example, if $p = 2$, then the sequence

$$1, 3, 7, 15, \ldots, 2^n - 1, \ldots$$

converges to -1, because the 2-adic distance between $2^n - 1$ and -1 is 2^{-n}, which goes to zero as $n \to \infty$. The metric is not complete: for example, you could prove by hand that the sequence of partial sums of the series

$$1 + p + p^4 + p^9 + \cdots + p^{n^2} + \cdots$$

are Cauchy but fail to converge to a limit in $\mathbb{Q}$, or later, after we have proved the Baire category theorem, we could use it to argue that a complete metric space containing $\mathbb{Q}$ in this metric must be uncountable. The completion of $\mathbb{Q}$ in this metric is called $\mathbb{Q}_p$, the p-adic numbers.

*This is unrelated to the metrics called d_p in Examples 1.5 and 1.8.

I will digress to discuss where the p-adic metric and its completion come from, but you can skip ahead if you like. They were developed by Hensel, who wanted to do something in number theory analogous to working on a power series "order by order." For example, we can study the equation $f(x)^2 = 1 - x$ by looking for a solution of the form

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots .$$

Looking at the constant terms, we find that a_0 could be 1 or -1. If we choose $a_0 = 1$, then looking at the linear terms we find that a_1 must be $-\frac{1}{2}$, looking at the quadratic terms we find that a_2 must be $-\frac{1}{8}$, and so on. (Later we could ask about the convergence of the power series that we obtain.) Hensel's idea is that if we want to study an equation like $x^2 = 2$, we can look for a solution that's a formal power series in a prime p, say $p = 7$,

$$x = a_0 + a_1 \cdot 7 + a_2 \cdot 7^2 + \cdots$$

with coefficients $a_i \in \{0, 1, \ldots, 6\}$. Again we can work order by order. Reducing modulo 7, we find that a_0 could be 3 or 4, because these are the two solutions to $a_0^2 \equiv 2 \pmod 7$. If we choose $a_0 = 3$, then reducing modulo $7^2 = 49$ we find that a_1 must be 1, reducing modulo $7^3 = 343$ we find that a_2 must be 2, and so on. The sequence of partial sums of the series

$$3 + 1 \cdot 7 + 2 \cdot 7^2 + \cdots$$

is $3, 10, 108, \ldots$ which is not Cauchy in the Euclidean metric – indeed, it goes to infinity. But it Cauchy is in the 7-adic metric, and the limit in the completion $\mathbb{Q}_p$ satisfies $x^2 = 2$. Questions of convergence for the power series correspond approximately to questions of whether the solution can be finagled back into $\mathbb{Q}$; this one cannot, because $\sqrt{2}$ is irrational.

An outstanding result in this area is the Hasse-Minkowski theorem, which states that a homogeneous quadratic equation in several variables, say $x^2 - xy + y^2 = 3z^2$, has a solution with coordinates in $\mathbb{Q}$ if and only if it has a solution in $\mathbb{R}$ and solutions in $\mathbb{Q}_p$ for every prime p.

By now we have gotten pretty far from real analysis or topology; the point is that while most of the examples we've considered are more or less closely related to the Euclidean metric on $\mathbb{R}$, the ideas we're developing have applications in radically different mathematical settings.

Example 3.9. The L^1 metric on $C([0,1])$ is not complete, as we can see using the sequence $f_1, f_2, f_3, \dots$ introduced at the beginning of the section: Exercise 3.1 asks you to prove that it is Cauchy, but let us sketch an argument that it does not converge to any limit $\ell \in C([0,1])$. We see that the restriction of f_n to an interval of the from $[0, \frac{1}{2} - \delta]$ converges to the constant function 0, and the restriction to $[\frac{1}{2} + \delta, 1]$ converges to the constant function 1. So if f_n converged to a limit ℓ, then ℓ would take the value 0 on the half-open interval $[0, \frac{1}{2})$, and 1 on $(\frac{1}{2}, 1]$, so it could not be continuous at $x = 1/2$.

As with $\mathbb{Q}$ in $\mathbb{R}$, one can embed $C([0,1])$ into a larger set of functions and extend the L^1 metric to a complete metric on that larger set, although this is far from staightforward. First, one has to to decide which discontinuous functions to include, which requires developing the Lebesgue integral. Second, the step function g that we want to be the limit is not uniquely determined at $x = 1/2$: we could set $g(1/2) = 0$, or $g(1/2) = 1$, or $g(1/2) = 1/2)$, or any other value, and $\int |f_n - g|$ will go to zero in any case, so one ends up working not with actual functions but with equivalence classes of functions. You can learn about all this in a course on measure theory.

While $C([0,1])$ is not complete in the L^1 metric, the next three propositions show that it is complete in the sup metric. We continue to assume the fact that every continuous function on $[0,1]$ is bounded.

Proposition 3.10. *Let (X,d) be a complete metric space, and let $Y \subset X$. Then Y is complete in the induced metric if and only if Y is closed in X.*

Proof. First suppose that Y is closed, and let $p_1, p_2, p_3, \dots$ be a Cauchy sequence in Y. By definition of the induced metric (Example 1.7), the sequence is also Cauchy in X, and because X is complete, it converges to a limit $\ell \in X$. Because Y is closed, we have $\ell \in Y$. Thus Y is complete.

Conversely, suppose that Y is complete, and let $p_1, p_2, p_3, \dots$ be a sequence in Y that converges to a limit $\ell \in X$. By Proposition 3.2, the sequence is Cauchy, so it also converges to a limit $\ell' \in Y$. By Exercise 1.9, we have $\ell = \ell'$, so $\ell \in Y$. Thus Y is closed. □

Proposition 3.11. *Let X be a set, let $B(X)$ be the set of bounded functions $f\colon X \to \mathbb{R}$, and define the sup metric on $B(X)$ by*

$$d(f,g) = \sup_{p \in X} |f(p) - g(p)|.$$

makes $B(X)$ into a complete metric space.

Proof. Let $f_1, f_2, f_3, \ldots \in B(X)$ be a sequence that is Cauchy in the sup metric. For each $p \in X$ we have $|f_m(p) - f_n(p)| \leq d(f_m, f_n)$, so the sequence $f_1(p), f_2(p), f_3(p), \ldots$ is also Cauchy in the usual metric on $\mathbb{R}$. Because $\mathbb{R}$ is complete, this sequence converges, so we can define a function $\ell \colon X \to \mathbb{R}$ by $\ell(p) = \lim_{n\to\infty} f_n(p)$.

Let us argue that $\ell \in B(X)$, that is, that ℓ is bounded. Again by Exercise 3.3, there is an $M \in \mathbb{R}$ such that $d(f_n, 0) = \sup |f_n| \leq M$ for all n, so $|f_n(p)| \leq M$ for all n and all $p \in X$. Letting $n \to \infty$, we get $|\ell(p)| \leq M$ for all $p \in X$, as desired.

Last we argue that $f_n \to \ell$ in the sup metric. Let $\epsilon > 0$ be given. Because the sequence $f_1, f_2, f_3, \ldots$ is Cauchy, there is an N such that if $m, n \geq N$ then $d(f_m, f_n) < \epsilon/2$, so $|f_m(p) - f_n(p)| < \epsilon/2$ for all $p \in X$. Letting $m \to \infty$, we get $|\ell(p) - f_n(p)| \leq \epsilon/2$ for all $n \geq N$ and all $p \in X$. Thus $d(\ell, f_n) \leq \epsilon/2 < \epsilon$. □

Proposition 3.12. *Continue to let $B(X)$ denote the set of bounded functions on a set X with the sup metric d. For any metric d_X on X, the set of bounded, continuous functions is closed in $B(X)$.*

Proof. Let $f_1, f_2, f_3, \ldots$ be a sequence of bounded, continuous functions converging in the sup metric to a bounded function ℓ. We want to prove that ℓ is continuous.

Let $p \in X$ and $\epsilon > 0$ be given. Because $f_n \to \ell$ in the sup metric, there is an N such that if $n \geq N$ then $\sup |f_n - \ell| < \epsilon/3$, and in particular $|f_N - \ell| < \epsilon/3$. Because f_N is continuous, there is a $\delta > 0$ such that if $d_X(p, q) < \delta$ then $|f_N(p) - f_N(q)| < \epsilon/3$. Thus if $d_X(p, q) < \delta$ then

$$
\begin{aligned}
&|\ell(p) - \ell(q)| \\
&\qquad \leq |\ell(p) - f_N(p)| + |f_N(p) - f_N(q)| + |f_N(q) - \ell(q)| \\
&\qquad\qquad < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon,
\end{aligned}
$$

where the first inequality is the triangle inequality in $\mathbb{R}$. □

Why completeness? You might ask why we care whether a metric space is complete. One major reason is the Banach fixed-point theorem, also called the contraction mapping theorem, which we prove next. It is crucial to proving the existence and uniqueness of solutions to ordinary differential equations, the inverse function theorem, and the implicit function theorem. Another reason is the Baire category theorem, to which we will return later.

Theorem 3.13 (Banach fixed-point theorem)**.** *Let (X, d) be a complete metric space, let $f : X \to X$, and suppose there is a "Lipschitz constant" $r \in [0, 1)$ such that for all $p, q \in X$ we have*

$$d(f(p), f(q)) \leq r \cdot d(p, q). \tag{3.2}$$

Then f has a unique fixed point, that is, there is a unique point $p \in X$ such that $f(p) = p$.

We can see that it is really necessary to have $r < 1$: if $X = \mathbb{R}$ with the usual metric, then a translation like $f(x) = x + 1$ satisfies (3.2) with $r = 1$, but it has no fixed point.

Proof of Theorem 3.13. Uniqueness is easy: if $f(p) = p$ and $f(q) = q$, then then

$$d(p, q) = d(f(p), f(q)) \leq r \cdot d(p, q),$$

so $d(p, q) = 0$, so $p = q$.

For existence, choose any point $p_0 \in X$, and define a sequence of points by repeatedly applying f:

$$p_1 = f(p_0) \qquad p_2 = f(p_1) \qquad p_3 = f(p_2) \qquad \ldots.$$

Let us argue that this sequence is Cauchy.

Set

$$D = d(p_0, p_1).$$

Then we have

$$d(p_1, p_2) = d(f(p_0), f(p_1)) \leq r \cdot d(p_0, p_1) = rD,$$

and

$$d(p_2, p_3) = d(f(p_1), f(p_2)) \leq r \cdot d(p_1, p_2) = r^2 D,$$

and similarly

$$d(p_n, p_{n+1}) \leq r^n D.$$

If we have some integer N and $n \geq m \geq N$, then by the triangle inequality,

$$\begin{aligned} d(p_m, p_n) &\leq d(p_m, p_{m+1}) + d(p_{m+1}, p_{m+2}) + \cdots + d(p_{n-1}, p_n) \\ &\leq r^m D + r^{m+1} D + \cdots + r^{n-1} D = \frac{r^m - r^n}{1 - r} \cdot D \leq \frac{r^N}{1 - r} \cdot D, \end{aligned}$$

where in the third step we have summed a geometric series.

Now let $\epsilon > 0$ be given. Because $r < 1$, we can choose an integer N such that

$$\frac{r^N}{1-r} \cdot D < \epsilon,$$

so if $m, n \geq N$ then $d(p_m, p_n) < \epsilon$. Thus the sequence is Cauchy, as claimed.

Because the metric is complete, there is an $\ell \in X$ such that $p_n \to \ell$. It remains to prove that $f(\ell) = \ell$. Exercise 3.4 asks you to prove that f is continuous. Thus it preserves limits by Exercise 1.11, and we can write

$$f(\ell) = f\left(\lim_{n\to\infty} p_n\right) = \lim_{n\to\infty} f(p_n) = \lim_{n\to\infty} p_{n+1} = \ell. \qquad \square$$

To give an idea of the power of the Banach fixed-point theorem, I'll digress to sketch the proof of existence of solutions to ordinary differential equations, but again you can skip ahead if you like. To take a concrete example, the motion of a pendulum in $\mathbb{R}^2$ is described by

$$x'(t) = y(t) \qquad\qquad y'(t) = -y(t) - \sin(x(t)),$$

which cannot be solved explicitly. Repackage this as follows: let $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ be given by $F(x, y) = (y, -y - \sin(x))$, and let $f\colon [0, 1] \to \mathbb{R}^2$ be the vector-valued function whose components are $x(t)$ and $y(t)$; then we want to solve $f'(t) = F(f(t))$ with some initial condition $f(0) = C \in \mathbb{R}^2$. (Any ordinary differential equation, including higher-order ones, can be packaged like this.)

Using the fundamental theorem of calculus, we can rewrite the differential equation $f'(t) = F(f(t))$ as an integral equation

$$f(t) = C + \int_0^t F(f(s))\, ds,$$

and observe that a solution is the same as a point $f \in C([0, 1])^2$ that's a fixed point of the map Φ from $C([0, 1])^2$ to itself given by $\Phi(f) = C + \int_0^t F(f(s))\, ds$. Using the fact that F is continuously differentiable, we can do some serious analysis and cook up a Lipschitz constant r for Φ as in (3.2); then the Banach fixed-point theorem gives a solution to our differential equation.

Moreover, the proof of the theorem yields a good algorithm for solving differential equations numerically: start with a constant function $f(t) = C$, and hit it with Φ over and over until you get close enough to a solution. You can learn about this in a course in differential equations or numerical methods, but for now let me just say that it is fun to see how applying this method to the differential equation $y' = y$ yields the power series for e^x.

Completion. We conclude this section by sketching how every metric space (X, d_X) can be embedded into a complete metric space $(\bar{X}, d_{\bar{X}})$. In examples, we usually have a more concrete way to construct the completion – we can construct $\mathbb{R}$ via Dedekind cuts, or $\mathbb{Q}_p$ via formal Laurent series in p, or $L^p([0,1])$ via Lebesgue integrable functions – but it is reassuring to know that a completion exists in any abstract situation.

We define $\bar{X}$ to be the set of equivalence classes of Cauchy sequences in X, where the equivalence relation is

$$\{p_n\} \sim \{p'_n\} \quad \Longleftrightarrow \quad \lim_{n\to\infty} d_X(p_n, p'_n) = 0. \tag{3.3}$$

Then we define

$$d_{\bar{X}}(\{p_n\}, \{q_n\}) = \lim_{n\to\infty} d_X(p_n, q_n). \tag{3.4}$$

There are many things to check:

(a) The relation (3.3) is an equivalence relation:

- Reflexive: $\{p_n\} \sim \{p_n\}$.
- Symmetric: if $\{p_n\} \sim \{p'_n\}$ then $\{p'_n\} \sim \{p_n\}$.
- Transitive: if $\{p_n\} \sim \{p'_n\}$ and $\{p'_n\} \sim \{p''_n\}$ then $\{p_n\} \sim \{p''_n\}$.

(b) The limit in (3.4) always exists.

(c) The function $d_{\bar{X}}$ defined in (3.4) is well-defined with respect to the equivalence relation (3.3): that is, if $\{p_n\} \sim \{p'_n\}$ and $\{q_n\} \sim \{q'_n\}$, then

$$d_{\bar{X}}(\{p_n\}, \{q_n\}) = d_{\bar{X}}(\{p'_n\}, \{q'_n\}).$$

(d) The function $d_{\bar{X}}$ defined in (3.4) is a metric, with the three properties given in Definition 1.2.

(e) The metric $d_{\bar{X}}$ on $\bar{X}$ is complete.

(f) While $\bar{X}$ does not literally contain X as a subset, the map $i\colon X \to \bar{X}$ that sends a point p to the constant sequence $p, p, p, \ldots$ preserves distances, so it identifies X with subset of $\bar{X}$.

(g) $\bar{X}$ is the smallest complete set that contains X, in the following sense: if (Y, d_Y) is a complete metric space and $j\colon X \to Y$ is a distance-preserving map, then there is a unique distance-preserving map $\bar{\jmath}\colon \bar{X} \to Y$ such that $\bar{\jmath} \circ i = j$.

Many of these are straightforward to prove: for example, for (c) we can write

$$\begin{aligned} d_{\bar{X}}(\{p_n\},\{q_n\}) &= \lim_{n\to\infty} d_X(p_n, q_n) \\ &\le \lim_{n\to\infty}\Big(d_X(p_n, p'_n) + d_X(p'_n, q'_n) + d_X(q_n, q'_n)\Big) \\ &= 0 + d_{\bar{X}}(\{p'_n\},\{q'_n\}) + 0, \end{aligned}$$

where the second step used the triangle inequality, and similarly

$$d_{\bar{X}}(\{p'_n\},\{q'_n\}) \le d_{\bar{X}}(\{p_n\},\{q_n\}),$$

so the two are equal. Exercise 3.5 asks you to prove (b). Here is a complete proof of (e), which is much harder than the others, but also terribly boring.

Proof of (e). Suppose we are given a Cauchy sequence in $\bar{X}$, that is, a Cauchy sequence of Cauchy sequences in X, where the second "Cauchy" is with respect to d_X and the first is with respect to $d_{\bar{X}}$. Let the first sequence be called $p_{1,1}, p_{1,2}, p_{1,3}, \ldots$, the second $p_{2,1}, p_{2,2}, p_{2,3}, \ldots$, and so on. We must construct a sequence $q_1, q_2, \ldots \in X$, prove that it is Cauchy with respect to d_X, and prove that our sequence of sequences converges to it with respect to $d_{\bar{X}}$

First let us define q_k. Because the sequence $p_{k,1}, p_{k,2}, p_{k,3}, \ldots$ is Cauchy with respect to d_X, we can apply the definition of Cauchy with $\epsilon = 1/k$ and get an integer N_k such that if $m, n \ge N_k$ then $d_X(p_{k,m}, p_{k,n}) < 1/k$. Then we set $q_k = p_{k,N_k}$.

Next let us argue that the sequence $q_1, q_2, q_3, \ldots$ is Cauchy with respect to d_X. Let $\epsilon > 0$ be given. Because the sequence $\{p_{1,n}\}, \{p_{2,n}\}, \{p_{3,n}\}, \ldots \in \bar{X}$ is Cauchy with respect to $d_{\bar{X}}$, there is an integer K such that if $k, l \ge K$ then

$$d_{\bar{X}}(p_{k,n}, p_{l,n}) < \epsilon/3. \tag{3.5}$$

Moreover we can increase K if necessary to get $1/K < \epsilon/3$. Fix some $k, l \ge K$. Recalling the definition of $d_{\bar{X}}$ from (3.4), we see that (3.5) says that there is an integer N such that if $n \ge N$ then

$$d_X(p_{k,n}, p_{l,n}) < \epsilon/3.$$

Set $n = \max(N_k, N_l, N)$. Then we have

$$\begin{aligned} d_X(q_k, q_l) &= d_X(p_{k,N_k},\, p_{l,N_l}) \\ &\le d_X(p_{k,N_k},\, p_{k,n}) + d_X(p_{k,n},\, p_{l,n}) + d_X(p_{l,n},\, p_{l,N_l}) \\ &< 1/k + \epsilon/3 + 1/l \le 1/K + \epsilon/3 + 1/K < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon, \end{aligned}$$

as desired.

Finally let us argue that $\{p_{1,n}\}, \{p_{2,n}\}, \{p_{3,n}\}, \ldots \to \{q_n\}$ with respect to $d_{\bar{X}}$. Let $\epsilon > 0$ be given. Because $\{q_n\}$ is Cauchy with respect to d_X, there is an integer K such that if $k, l \geq K$ then $d_X(q_k, q_l) < \epsilon/2$. Increase K if necessary to get $1/K < \epsilon/2$. If $k \geq K$ then

$$d_{\bar{X}}(\{p_{k,n}\}, \{q_n\}) = \lim_{n\to\infty} d_X(p_{k,n}, q_n) \leq \lim_{n\to\infty} d_X(p_{k,n}, q_k) + \lim_{n\to\infty} d_X(q_k, q_n)$$

The first term is

$$\lim_{n\to\infty} d_X(p_{k,n}, p_{k,N_k}),$$

where N_k is the one we chose in the previous paragraph; by construction, if $n \geq N_k$ then $d_X(p_{k,n}, p_{k,N_k}) < 1/k$, so the limit is $\leq 1/k < \epsilon/2$. For the second term, if $n \geq K$ then $d_X(q_k, q_n) < \epsilon/2$, so the limit is $\leq \epsilon/2$. Thus the sum is $< \epsilon$, as desired. $\square$

Exercises.

3.1. Prove that the sequence of piecewise-linear functions $f_1, f_2, f_3, \ldots \in C([0,1])$ introduced at the beginning of the section is Cauchy in the L^1 metric.

3.2. Let $X = \mathbb{Q}$, let p be a prime number, and let d_p be the p-adic meric defined in Example 3.8.

(a) Write down some rational numbers, and compute the 2-adic distance between them.

(b) Prove that d_p is a metric.

(c) Consider a "formal power series" in p,

$$a_0 + a_1 p + a_2 p^2 + a_3 p^3 + \cdots,$$

where the coefficients a_n are integers with $0 \leq a_n < p$. Let $s_0, s_1, s_2, \ldots \in \mathbb{Z}$ be the sequence of partial sums of this series. Prove that it is Cauchy in the p-adic metric.

3.3. Let $p_1, p_2, p_3, \ldots$ be a Cauchy sequence in a metric space (X, d). Prove that the sequence is bounded, meaning that there is a point $q \in X$ and a radius $R > 0$ such that $p_n \in B_R(q)$ for all n. In fact, for any $q \in X$ you can find such a radius R, and in particular for $X = \mathbb{R}$ you can take $q = 0$.

Hint: Start by applying the definition of Cauchy with $\epsilon = 1$ to get an N such that if $m, n \geq N$ then $d(p_m, p_n) < 1$, and in particular $d(p_N, p_n) < 1$.

3.4. Let (X, d) be a metric space, let $f\colon X \to X$, suppose there is a "Lipschitz constant" $r \in [0, 1)$ such that for all $p, q \in X$ we have

$$d(f(p), f(q)) \leq r \cdot d(p, q).$$

Prove that f is continuous.

Hint: Take $\delta = \epsilon$.

3.5. Prove item (b) from the laundry list of things to check about the completion of a metric space: if $\{p_n\}$ and $\{q_n\}$ are Cauchy sequences in X, then the limit

$$\lim_{n\to\infty} d_X(p_n, q_n)$$

exists.

Hint: Prove that $d_X(p_n, q_n)$ is a Cauchy sequence in $\mathbb{R}$ with its usual metric. The reverse triangle inequality (Exercise 1.5) may be useful.

4 Interior, Closure, and Boundary

Let (X, d) be a metric space, and consider a subset $A \subset X$.

Definition 4.1. The *interior* of A, denoted $\operatorname{int} A$ or sometimes A°, is

$$\{p \in X : B_r(p) \subset A \text{ for some } r > 0\}.$$

We see that $\operatorname{int} A \subset A$, and the two are equal if and only if A is open.

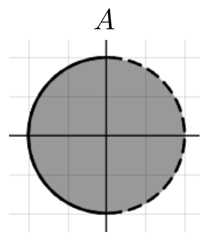
Definition 4.2. The *closure* of A, denoted $\bar{A}$ or sometimes $\operatorname{cl} A$, is

$$\{p \in X : \text{there is a sequence } a_1, a_2, a_3, \ldots \in A \text{ that converges to } p\}.$$

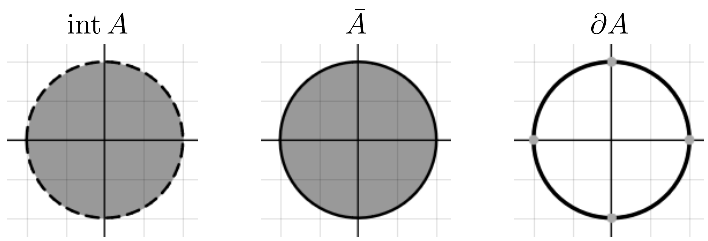
We see that $A \subset \bar{A}$, and the two are equal if and only if A is closed.

Definition 4.3. The *boundary* of A, denoted ∂A, is $\bar{A} \setminus \operatorname{int} A$.

Example 4.4. Let $X = \mathbb{R}^2$ with the Euclidean metric, and let A be the half-open disc shown below:

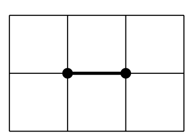


The interior is the open disc $x^2 + y^2 < 1$. The closure is the closed disc $x^2 + y^2 \leq 1$. The boundary is the unit circle $x^2 + y^2 = 1$.



Example 4.5. The interior, closure, and boundary very much depend on the ambient space: if we make the ambient X bigger or smaller, then they will change. For example, if we take $A = [0, 1]$ inside $\mathbb{R}$ with the usual metric,

then the interior is the open interval $(0, 1)$, the closure is the same as A, and the boundary is the two points 0 and 1. But if we take it in $\mathbb{R}^2$ with the Euclidean metric as shown,



then the interior is empty, the closure is the same as A, and the boundary is also the same as A.

Example 4.6. Let $X = \mathbb{R}$ with the usual metric, and let $A = \mathbb{Q}$. Then the interior of A is empty, the closure is all of $\mathbb{R}$, and the boundary is all of $\mathbb{R}$.

Here is another characterization of the closure:

Proposition 4.7. *Let (X, d) be a metric space, and let $A \subset X$.*

$$\bar{A} = \{p \in X : B_r(p) \cap A \neq \varnothing \text{ for all } r > 0\}.$$

Proof. First suppose that $p \in \bar{A}$, and let $a_1, a_2, a_3, \ldots$ be a sequence in A that converges to p. Then for every $r > 0$ there is an N such that if $n \geq N$ then $d(a_n, p) < r$. Thus $B_r(p) \cap A$ contains a_N, so it is not empty.

Conversely, suppose that $B_r(p) \cap A$ is not empty for all $r > 0$. For $n = 1, 2, 3, \ldots$, take $r = 1/n$, and choose a point $a_n \in B_{1/n}(p) \cap A$. Then the sequence $a_1, a_2, a_3, \ldots$ converges to p so $p \in \bar{A}$. □

Proposition 4.8. *The complement of the interior is the closure of the complement, and vice versa: that is, if (X, d) is a metric space and $A \subset X$, then*

$$X \setminus \operatorname{int} A = \overline{X \setminus A} \quad \text{and} \quad X \setminus \bar{A} = \operatorname{int}(X \setminus A).$$

Proof. The two claims are equivalent: replacing A with $X \setminus A$ turns one into the other. To prove the first equality, observe that each of the following statements is equivalent to the next:

- $p \in X \setminus \operatorname{int} A$.
- $p \notin A$.
- $B_r(p) \not\subset A$ for all $r > 0$.
- $B_r(p) \cap (X \setminus A) \neq 0$ for all $r > 0$.
- $p \in \overline{X \setminus A}$.

Here the last step used Proposition 4.7. □

Corollary 4.9. *Let (X, d) be a metric space and $A \subset X$. Then*

$$\partial A = \partial(X \setminus A) = \bar{A} \cap \overline{X \setminus A}. \qquad \square$$

Here are some basic properties of interiors. Exercise 4.2 asks you to prove the analogous properties of closures.

Proposition 4.10. *Let (X, d) be a metric space, and let $A, B \subset X$.*

(a) $\operatorname{int} A$ *is open.*

(b) $\operatorname{int} A$ *is the biggest open set contained in* A*, in the following sense: if* $U \subset X$ *is open and* $U \subset A$*, then* $U \subset \operatorname{int} A$*.*

(c) If $A \subset B$ *then* $\operatorname{int} A \subset \operatorname{int} B$*.*

(d) $\operatorname{int}(A \cap B) = \operatorname{int} A \cap \operatorname{int} B$*.*

(e) $\operatorname{int} A \cup \operatorname{int} B \subset \operatorname{int}(A \cup B)$*, but Example 4.11 below shows that the inclusion can be strict.*

Proof. (a) By definition, for every $p \in \operatorname{int} A$ there is an $r > 0$ such that $B_r(p) \subset A$. Let us argue that $B_r(p) \subset \operatorname{int} A$, which will prove that $\operatorname{int} A$ is open. This is very similar to the proof that an open ball is open in Exercise 2.5(a).

So we need to prove that for every $q \in B_r(p)$, there is an $s > 0$ such that $B_s(q) \subset A$. Take $s = r - d(p,q)$, which is positive because $d(p,q) < r$. If $q' \in B_s(Q)$ then

$$d(p,q') \leq d(p,q) + d(q,q') < d(p,q) + s = r,$$

so $q' \in B_r(p)$, so $q' \in A$.

(b) Let $p \in U$. Because U is open, then there is an $r > 0$ such that $B_r(p) \subset U$. Because $U \subset A$, we have $B_r(p) \subset A$. Thus $p \in \operatorname{int} A$.

(c) We have $\operatorname{int} A \subset A \subset B$, and $\operatorname{int} A$ is open by part (a), so $\operatorname{int} A \subset \operatorname{int} B$ by part (b).

(d) We have $A \cap B \subset A$, so $\operatorname{int}(A \cap B) \subset \operatorname{int} A$ by part (c), and similarly $\operatorname{int}(A \cap B) \subset \operatorname{int} B$; thus $\operatorname{int}(A \cap B) \subset \operatorname{int} A \cap \operatorname{int} B$.

For the reverse inclusion, we have $\operatorname{int} A \cap \operatorname{int} B \subset \operatorname{int} A \subset A$, and $\operatorname{int} A \cap \operatorname{int} B \subset \operatorname{int} B \subset B$, so $\operatorname{int} A \cap \operatorname{int} B \subset A \cap B$. But $\operatorname{int} A \cap \operatorname{int} B$ is an intersection of two open sets by part (a), hence is open by Exercise 2.6(a), so $\operatorname{int} A \cap \operatorname{int} B \subset \operatorname{int}(A \cap B)$ by part (b).

(e) We have $\operatorname{int} A \subset A \subset A \cup B$, and $\operatorname{int} B \subset B \subset A \cup B$, so $\operatorname{int} A \cup \operatorname{int} B \subset A \cup B$. Thus $\operatorname{int} A \cup \operatorname{int} B \subset \operatorname{int}(A \cup B)$ by part (c). □

Example 4.11. It need not be true that $\operatorname{int} A \cup \operatorname{int} B = \operatorname{int}(A \cup B)$. For example, let $X = \mathbb{R}$ with the usual metric, let $A = [0,1]$, and let $B = [1,2]$, so $A \cup B = [0,2]$. Then $\operatorname{int}(A \cup B) = (0,2)$ is bigger than $\operatorname{int} A \cup \operatorname{int} B = (0,1) \cup (1,2)$.

Exercises.

4.1. Find the closure, interior, and boundary of each subset of $\mathbb{R}^2$ in the Euclidean topology:

(a) $A_1 = \{(x, y) : 0 < x \leq 1, 0 \leq y < 1\}$

(b) $A_2 = \{(x, y) : 0 < x \leq 1, y = 0\}$

(c) $A_3 = \{(x, y) : x \in \mathbb{Q} \text{ or } y \in \mathbb{Q}\}$

(d) The subset A from Exercise 2.1.

4.2. Prove the analogue of Proposition 4.10 for closures without using Proposition 4.8. For (c), (d), and (e) especially, you'll want to follow the proof Proposition 4.10 closely.

(a) $\bar{A}$ is closed.
Hint: It's easier to use Proposition 4.7.

(b) $\bar{A}$ is the smallest closed set contained in A, in the following sense: if $F \subset X$ is closed and $A \subset F$, then $\bar{A} \subset F$.

(c) If $A \subset B$ then $\bar{A} \subset \bar{B}$.

(d) $\overline{A \cup B} = \bar{A} \cup \bar{B}$.

(e) $\overline{A \cap B} \subset \bar{A} \cap \bar{B}$. Give an example to show that the inclusion can be strict.

4.3. Exercise 2.5 asked you to prove that the open ball $B_r(p)$ is open, and the closed ball $\bar{B}_r(p)$ is closed. Thus $B_r(p) \subset \operatorname{int} \bar{B}_r(p)$ by Proposition 4.10(b), and $\overline{B_r(p)} \subset \bar{B}_r(p)$ by Exercise 4.2(b). But give an example to show that the inclusions can be strict.

Hint: You might take $X = \mathbb{Z}$ with the usual metric inherited from $\mathbb{R}$, or any set with a discrete metric (Exercise 1.8).

References

[1] Douglas Dunham. M. C. Echer's use of Poincaré models of hyperbolic geometry. Available at `www.math-art.eu/Documents/pdfs/Dunham.pdf`.

[2] Felix Hausdorff. *Grundzüge der Mengenlehre.* Verlag von Veit & Comp, Leipzig, 1914.

[3] Heinrich Tietze. Beiträge zur allgemeinen Topologie. I. *Math. Ann.*, 88(3-4):290–312, 1923.