Lecture Notes - Real Analysis

MA Math Camp 2022
Columbia University

Andrea Ciccarone*

Last Updated : August 15, 2022

Contents

1 Metric Spaces and Topology **2**
1.1 Distance 2
1.2 Open Balls 3
1.3 Open sets and closed sets 4

2 Sequences and Convergence **7**
2.1 General Definitions 7
2.2 Convergence in Euclidean Spaces 8
2.3 Monotone Convergence and Bolzano-Weierstrass Theorem 10

3 Compactness **11**
3.1 General Properties 11
3.2 Heine-Borel Theorem in $\left(\mathbb{R}^k, d_2\right)$ 12
3.3 Sequential Compactness 14

4 Cauchy Sequences and Completeness **14**
4.1 Contraction Mapping Theorem 16

5 Continuity of Functions **18**
5.1 Limits of Functions 18
5.2 Continuity 19
5.3 Weierstrass Theorem 21

6 Appendix **23**

*The present lecture notes were largely based on math camp materials from César Barilla, Palaash Bhargava, Paul Koh, and Xuan Li. All errors in this document are mine. If you find a typo or an error, please send me an email at ac4790@columbia.edu.

1 Metric Spaces and Topology

1.1 Distance

In Calculus, the distance $d(a, b)$ between two points a and b of the real line $\mathbb{R}$ is given by $|a-b|$, while for two vectors x and y of $\mathbb{R}^n$ their distance $d(x, y)$ is given by $\sqrt{\sum_{i=1}^n (x_i - y_i)^2}$. Our aim in this chapter is to extend the notion of distance to abstract spaces. To this end, the first thing to observe is that the distance among vectors just mentioned is certainly not the only possible one. For example, consider two vectors $x = (x_1, x_2)$ and $y = (y_1, y_2)$ in the plane $\mathbb{R}^2$. Suppose that these two vectors represent the coordinates of two places in a town with a square plan. The length of the shortest way for a pedestrian to move between them is certainly not given by $\sqrt{\sum_{i=1}^n (x_i - y_i)^2}$, that is, by the length of the segment that joins the points x and y (a segment that could be covered only by an hypothetical subway joining the two places). Looking at the map, it is easy to see that the effective distance is given by $|x_1 - y_1| + |x_2 - y_2|$. Formally, given two vectors $x, y \in \mathbb{R}^n$, we define the distance $d(x, y)$ as $\sum_{i=1}^n |x_i - y_i|$. In the case $n = 2$ we find again the "pedestrian" distance $d(x, y) = |x_1 - y_1| + |x_2 - y_2|$ just discussed.

To see another example of distance, suppose that two vectors x and y in $\mathbb{R}^n$ denote the allocations of income in a society composed by n individuals. Therefore, x_i is the income that individual i has under allocation x, while y_i is his income under allocation y. How we measure the the "distance" between the allocations x and y? A possibility is to evaluate the individual differences of income $|x_i - y_i|$ among the two allocations, and to take the quantity $\max_{1 \le i \le n} |x_i - y_i|$ as the distance between the two allocations x and y. In other words, we evaluate the distance between the two allocations by considering the individual whose income is subject to the greatest variation (in absolute value). Given two vectors $x, y \in \mathbb{R}^n$, we define therefore the distance $d(x, y) = \max_{1 \le i \le n} |x_i - y_i|$

We try now to abstract from the particular examples, in order to arrive at a general denition of distance. We observe that the distances just discussed have the following properties:

- The distance between two vectors is always non-negative: $d(x, y) \ge 0$
- Two vectors have zero distance if and only if they coincide: $d(x, y) = 0$ if and only if $x = y$
- The distance between two vectors is symmetric: $d(x, y) = d(y, x)$
- Given three vectors, the triangular inequality holds: $d(x, y) \le d(x, z) + d(z, y)$

All this leads us to the following definition.

Definition 1.1. *Let X be a set. A function $d : X^2 \to \mathbb{R}_+$ is a* ***distance function****, or* ***metric****, on X iff it satisfies the following properties*

(1) $d(x, y) = 0$ iff $x = y$,

(2) Symmetry: $d(x, y) = d(y, x)$, $\forall\ x, y \in X$, and

(3) Triangle inequality: $d(x, y) \le d(x, z) + d(z, y)$, $\forall\ x, y, z \in X$

If d is a metric on X, then the couple (X, d) is called a ***metric space****.*

Elements of a metric space are often called **points**. Note that the distance function d is an inseparable part of a metric space. (X, d_1) and (X, d_2) are two different metric spaces if d_1 and d_2 are two different distance functions. We can denote a metric space simply as X only when there is no ambiguity regarding what distance function is used.

As an example, the set $\mathbb{R}^k$ can be endowed with a natural metric, the **Euclidean distance function** d_2, defined as

$$d_2(x,y) := \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2}$$

for any $x, y \in \mathbb{R}^k$. This distance function is natural in the sense that it is consistent with how we understand "distance" in our real life where $k = 3$. Observe that when $k = 1$, $d_2(x,y) = |x - y|$ i.e we recover the absolute value as our primitive notion of distance in dimension one. We can easily verify that d_2 satisfies properties (1) (2) (3)[1]. The metric space $\left(\mathbb{R}^k, d_2\right)$ is often called a Euclidean space.

Examples of other metrics on $\mathbb{R}^k$:

- d_n metric:

$$d_n(x,y) := \left(\sum_{i=1}^{k}|x_i - y_i|^n\right)^{\frac{1}{n}}$$

  where k can be any positive integer[2]. This subsumes the Euclidean distance d_2 as a special case. Notice that when $k = 1$, all d_n's reduce to the same absolute distance function $d(x,y) := |x - y|$.

- d_∞ metric :

$$d_\infty(x,y) = \sup_{j=1,\dots,k}|x_k - y_k|$$

  This extends the d_n metric to the case $n = \infty$.

- discrete metric:

$$d(x,y) := \begin{cases} 0, \text{ if } x = y \\ 1, \text{ if } x \neq y \end{cases}$$

Given a metric space (X, d) and a subset $S \subset X$, we can restrict ourselves to the subset S to get a smaller metric space, still using the distance function defined on the larger space X. Formally, define a new distance function $d|_S : S^2 \to \mathbb{R}_+$ as $d|_S(x,y) := d(x,y)$ for any $x, y \in S$. That is, $d|_S$ is the original distance function d restricted in the subset S. Clearly, $d|_S$ is a valid metric on S, and so $\left(S, d|_S\right)$ is a valid metric space. Sometimes, we write the metric space $\left(S, d|_S\right)$ as (S, d) for simplicity, but keep in mind that rigorously speaking the new metric d in (S, d) has a different domain from the original metric.

1.2 Open Balls

Definition 1.2. *Let (X, d) be a metric space. The **open ball** centered at $x \in X$ with radius $r > 0$ is defined as the set*

$$B_r(x) := \left\{z \in X : d(z,x) < r\right\}$$

[1] The triangle inequality can be shown using Cauchy-Schwarz inequality.
[2] The triangle inequality can be shown using Minkowski inequality.

Keep in mind that an open ball $B_r(x)$ depends both on the whole space X and the distance function d. For example, in the metric space $(\mathbb{R}, d_2)$, the open ball $B_1(0) = (-1, 1)$; however, in $(\mathbb{R}_+, d_2)$, the open ball $B_1(0) = [0, 1)$. If we use the discrete metric d, then in $(\mathbb{R}, d)$ the open ball $B_1(0) = \{0\}$. Therefore, when we write down the notation for an open ball like $B_1(0)$, we have to be clear about which metric space we are working with.

The notations $B_r(x)$ or $B(x, r)$ are sometimes used alternatively to denote the open ball of radius r centered at x.

Definition 1.3. *Let (X, d) be a metric space and S be a subset of X. The set S is said to be **bounded** iff there exists $x \in X$ and $r > 0$ s.t. $B_r(x) \supset S$.*

That is, a set S is bounded iff we can bound it using an open ball – in other words, we can contain the whole set in that ball. This captures the idea that there is a finite maximal distance between any two points in that set.

1.3 Open sets and closed sets

Definition 1.4. *Let (X, d) be a metric space, and S a subset of X.*

*A point $x \in X$ is an **interior point** of S iff $\exists\, r > 0$ s.t. $B_r(x) \subset S$. The set of interior points of S is denoted as $int(S)$.*

*The set S is an **open set** iff $S \subset int(S)$, i.e. all points in S are interior points.*

Clearly, any interior point of S is a point in S, and therefore, $S \subset int(S)$ is equivalent to $S = int(S)$.

By convention, the empty set $\emptyset$ is open, because $int\left(\emptyset\right) = \emptyset$. Also, the whole space X is also open, because $int(X) = X$. Keep in mind that whether a set is open depends on the metric space. For example, $[0, 1)$ is not an open set in $(\mathbb{R}, d_2)$, since 0 is not an interior point. However, $[0, 1)$ is an open set in $(\mathbb{R}_+, d_2)$. The point 0 becomes an interior point of $[0, 1)$ because an open ball centered at 0 is now $B_r(0) = [0, r)$ instead of $(-r, r)$.

As its name suggests, an open ball is open.

Claim 1.5. *In metric space (X, d), any open ball is an open set.*

Proof. Take any open ball $B_r(x)$ in the metric space. In order to show that the open ball $B_r(x)$ is an open set, we need to show for any $z \in B_r(x)$, $z \in int(B_r(x))$, i.e. $\exists \epsilon$ s.t. $B_\epsilon(z) \subset B_r(x)$.

Take any point $z \in B_r(x)$. Let $\varepsilon := r - d(z, x)$.

First, because $z \in B_r(x)$, we have $d(z, x) < r$ and thus $\varepsilon > 0$.

Second, take any $y \in B_\varepsilon(z)$, we have

$$d(y, x) \le d(y, z) + d(z, x) < \varepsilon + d(z, x) = r$$

and therefore $y \in B_r(x)$. □

Note that open intervals are open in $(\mathbb{R}, d_2)$, because they are special cases of open balls.

The next proposition is an important property of open sets. It states that an arbitrary union of open sets is open, and that a finite intersection of open sets is also open.

Proposition 1.6. *In metric space (X, d):*

(1) Let $\{E_\alpha\}_{\alpha \in A}$ be an arbitrary family of open sets (potentially uncountably many of them). Then their union $\bigcup_{\alpha \in A} E_\alpha$ is also open.

(2) Let $\{E_i\}_{i=1}^n$ be a finite family of open sets. Then their intersection $\bigcap_{i=1}^{n} E_i$ is also open.

Proof. (1) Take any $x \in \bigcup_{\alpha \in A} E_\alpha$, we need to find $r > 0$ s.t. $B_r(x) \subset \bigcup_{\alpha \in A} E_\alpha$.

By definition of union, $\exists\ \hat{\alpha} \in A$ s.t. $x \in E_{\hat{\alpha}}$. Because $E_{\hat{\alpha}}$ is open, we can find $r > 0$ s.t. $B_r(x) \subset E_{\hat{\alpha}}$.

This is an r we need to find because $B_r(x) \subset E_{\hat{\alpha}} \subset \bigcup_{\alpha \in A} E_\alpha$.

(2) Take any $x \in \bigcap_{i=1}^{n} E_i$, we need to find $r > 0$ s.t. $B_r(x) \subset \bigcap_{i=1}^{n} E_i$.

By definition of intersection, $x \in E_i$ for any $i = 1, 2, \ldots, n$. For each i, because E_i is open, $\exists\ r_i > 0$ s.t. $B_{r_i}(x) \subset E_i$.

Let $r := \min\{r_1, r_2, \ldots, r_n\}$, and this is an r we need to find. First, clearly $r > 0$. Second, $B_r(x) \subset B_{r_i}(x) \subset E_i$ for any i, and therefore $B_r(x) \subset \bigcap_{i=1}^{n} E_i$.

$\square$

Note that an infinite intersection of open sets may not be open. For example, consider $E_n = \big(-1/n, 1/n\big)$, and we have $\bigcap_{n=1}^{+\infty} E_n = \{0\}$.

Now let's move on to closed sets. We give two equivalent definitions.

Definition 1.7. *Let (X, d) be a metric space, and S a subset of X; denote by $S^c := X \setminus S$ its complement. We say that S is closed is S^c is open.*

Remark 1.8. *In this lecture, we consider only* metric topology *: we first a metric space, and this induces a topology (i.e open and closed sets) through the metric we chose. This is not the only way to define a topology. In general abstract topology, the approach is actually diametrically reversed : we start from the fundamental data of the class of open sets (which sets are defined as open) and build from there. This is where the canonical (and probably simplest) definition of closed sets as the complements of open sets comes from.*

A corollary of is that S is an open set iff S^c is a closed set. Simply put, the complement of an open set is closed, and the complement of a closed set is open.The next definition of closed sets is equivalent to the previous one (in the context of metric spaces).

Definition 1.9. *Let (X, d) be a metric space, and S a subset of X.*

A point $x \in X$ is a ***limit point*** *of S iff $\big(B_r(x) \setminus \{x\}\big) \cap S \neq \emptyset$, $\forall\ r > 0$. The set of limit points of S is denoted as S'.*

The set S is a ***closed set*** *iff $S \supset S'$, i.e. S contains all of its limit points.*

The condition "$\big(B_r(x) \setminus \{x\}\big) \cap S \neq \emptyset$, $\forall\ r > 0$" states that the open ball $B_r(x)$ with the center removed always contains some points in the set S, no matter how small the radius r is. That is, a point x is a limit point of S iff we can use points in S to approximate x arbitrarily well (the point x itself may be a point in S, but we are not allowed to use x to approximate itself).

Notice that not every point in S is necessarily a limit point, and so $S \supset S'$ is *not* equivalent to $S = S'$. For example in $(\mathbb{R}, d_2)$, the "isolated" point 2 in the set $S = [0, 1] \cup \{2\}$ is not a limit point of S. The set S is indeed closed by definition, since $S' = [0, 1]$ which is a proper subset of S.

Exercise 1.10. *Prove that the two definitions above for closed sets are equivalent.*

By convention, the empty set $\emptyset$ is closed, because $\emptyset' = \emptyset$. Also, the whole space X is also closed, because a limit point of X, by definition, must be a point in X in the first place. Keep in mind that whether a set is closed depends on the metric space. For example, $(0, 1]$ is not a closed set in $(\mathbb{R}, d_2)$, since it does not contain its limit point 0. However, $(0, 1]$ is a closed set in $(\mathbb{R}_{++}, d_2)$. The point 0 is no longer a limit point of $(0, 1]$, since it is not even a point because it is not in the metric space.

The following proposition establishes a third characterization of closed sets : the *sequential characterization* (this anticipates slightly on the next section, but we have already seen sequences in the previous lecture).

Proposition 1.11. *Let (X, d) be a metric space, and S a subset of X. Then the following two statements are equivalent.*

(1) S is a closed set.

(2) (sequential definition) For any sequence (x_n) in S convergent to some point $x \in X$, we have $x \in S$.

Proof. $\Rightarrow$:

Take any sequence (x_n) in S convergent to $x \in X$, WTS $x \in S$.

Suppose that $x \notin S$, WTS $x \in S'$.

Take any $r > 0$. Because $x_n \to x$, $\exists\ n$ s.t. $x_n \in B_r(x)$. Because $x_n \in S$ but $x \notin S$, we know that $x_n \neq x$, and thus $x_n \in B_r(x) \setminus \{x\}$. Therefore, $x_n \in \big(B_r(x) \setminus \{x\}\big) \cap S$, which implies $\big(B_r(x) \setminus \{x\}\big) \cap S \neq \emptyset$.

So we have shown that $x \in S'$.

Because S is closed, we have $x \in S' \subset S$, which contradicts the hypothesis $x \notin S$ we started with.

Therefore, it must be the case that $x \in S$.

$\Leftarrow$:

Take any $x \in S^c$. We want to find $r > 0$ s.t. $B_r(x) \subset S^c$.

Suppose that we cannot find such r, then for any $n \in \mathbb{N}$, we have $B_{1/n}(x) \not\subset S^c$. Then for each n, there exists $x_n \in B_{1/n}(x)$ s.t. $x_n \in S$. Clearly, $x_n \to x$ because for any $\varepsilon > 0$, we can let N be some number greater than $1/\varepsilon$, and then for any $n > N$, we have $d(x_n, x) < 1/n < 1/N < \varepsilon$.

Hence we have $x \in S$, which contradicts $x \in S^c$. □

The next proposition is an important property of closed sets. It states that an arbitrary intersection of closed sets is closed, and that a finite union of closed sets is also closed.

Proposition 1.12. *In metric space (X, d):*

(1) Let $\{F_\alpha\}_{\alpha \in A}$ be an arbitrary family of closed sets (potentially uncountably many of them). Then their intersection $\bigcap_{\alpha \in A} F_\alpha$ is also closed.

(2) Let $\{F_i\}_{i=1}^n$ be a finite family of closed sets. Then their union $\bigcup_{i=1}^n F_i$ is also closed.

The proposition above is simply a corollary of Proposition 1.6, using De Morgan's law and the fact that the complement of an open set is closed. This is left as an exercise.

As a final note, again keep in mind that open sets and closed sets are not "absolute" concepts. They rely on the metric space we are working with. When there is ambiguity regarding which metric space we are using, we have to be explicit about it by saying "set S is open/closed in the metric space (X,d)" instead of simply saying "S is open/closed". Also, notice that under discrete metric, all sets in the metric space are both open and closed (exercise).

Please use the following examples to check your understanding of open sets and closed sets.

	$[0,+\infty)$	$(0,+\infty)$	$\{1/n : n \in \mathbb{N}\}$
In $(\mathbb{R}, d_2)$	not open, but closed	open, not closed	not open, not closed
In $(\mathbb{R}_+, d_2)$	open and closed	open, not closed	not open, not closed
In $(\mathbb{R}_{++}, d_2)$	NA	open and closed	not open, but closed

2 Sequences and Convergence

2.1 General Definitions

Definition 2.1. *Let X be a set. The function $x : \mathbb{N} \to X$ is called a **sequence** in X.*

Sequences are simply a special case of functions. The value of the function x evaluated at 1, $x(1)$, is called the first **term** of the sequence, and the value of the function evaluated at 2, $x(2)$, is called the second term, and so on. By convention, we often use subscripts and write $x_1, x_2, \ldots$ instead of $x(1), x(2), \ldots$, and the whole sequence is often denoted as (x_n) instead of x.

Note that there is no distance function involved in the definition above, since we don't need a concept of distance to talk about sequences. However, we do need distance to talk about convergence.

Definition 2.2. *Let (X,d) be a metric space. A sequence (x_n) in X is said to **converge** to a point $x \in X$, iff $\forall\, \varepsilon > 0$, $\exists\, N \in \mathbb{N}$ s.t. $d(x_n, x) < \varepsilon$ for all $n > N$.*

*When the sequence (x_n) converges to x, the point x is called a **limit** of the sequence (x_n), and we use the notation $x_n \to x$ or $\lim_{n\to\infty} x_n = x$.*

Notice that the requirement $d(x_n, x) < \varepsilon$ is equivalent to $x_n \in B_\varepsilon(x)$. Another way to describe convergence is that the sequence (x_n) will eventually go into the open ball $B_\varepsilon(x)$, no matter how small the ball is.

The next claim establishes that the limit of a convergent sequence must be unique, and therefore it makes sense to talk about "the" limit of a convergent sequence.

Claim 2.3. *Let (X,d) be a metric space. Suppose $x_n \to x$ and $x_n \to x'$, then $x = x'$.*

Proof. We prove this claim by contradiction.

Suppose $x \neq x'$. By property (1) of d, we have $d(x, x') > 0$.

Let $\varepsilon := d(x, x')/2$. Because $x_n \to x$, there exists N s.t. $d(x_n, x) < \varepsilon$ for any $n > N$. Because $x_n \to x'$, there exists N' s.t. $d(x_n, x') < \varepsilon$ for any $n > N'$. Let $\hat{n} := \max\{N, N'\} + 1$, and so we have $\hat{n} > N$ and $\hat{n} > N'$. Therefore, we have $d(x_{\hat{n}}, x) < \varepsilon$ and $d(x_{\hat{n}}, x') < \varepsilon$, and thus

$$d(x_{\hat{n}}, x) + d(x_{\hat{n}}, x') < 2\varepsilon = d(x, x')$$

which contradicts triangle inequality of d.

Therefore we must have $x = x'$. □

A sequence is said to be **bounded** iff its range $\{x_1, x_2, \ldots\}$ is a bounded set. The next claim establishes that a convergent sequence must be bounded.

Claim 2.4. *Let (X, d) be a metric space. If (x_n) is a convergent sequence in X, then (x_n) must be bounded.*

Proof. Let the limit of (x_n) be x. Let $\varepsilon = 1$, and by definition of convergence, there exists N s.t. $d(x_n, x) < 1$ for any $n > N$. Then let

$$r := \max\left\{d(x_1, x), d(x_2, x), \ldots, d(x_N, x)\right\} + 1$$

Then clearly we have $B_r(x) \supset \{x_1, x_2, \ldots\}$. $\square$

The trick of this proof is to cut the sequence into a "head" and a "tail". Then we use convergence to bound the tail, and the head is automatically bounded because it has finitely many terms.

The discussion above applies to general metric spaces. The next subsection is dedicated to an important special case, the Euclidean spaces, and establishes some more results on convergence in Euclidean spaces.

2.2 Convergence in Euclidean Spaces

The next claim states that in $\mathbb{R}$ the $\leq$ relation is preserved in the limit.

Claim 2.5. *In $(\mathbb{R}, d_2)$, let there be two convergent sequences $x_n \to x$ and $y_n \to y$. If $x_n \leq y_n$ for any $n \in \mathbb{N}$, then $x \leq y$.*

We can prove this claim by contradiction, and this is left as an exercise. Also note that this is not true if we replace $\leq$ by $<$, since for example, $x_n = -1/n$ and $y_n = 1/n$.

The next proposition states that a sequence of vectors converges to a limit vector iff each coordinate converges separately.

Proposition 2.6. *Let (x_n) be a sequence in $\left(\mathbb{R}^k, d_2\right)$. The sequence (x_n) converges to $x \in \mathbb{R}^k$ iff the sequence $\left(x_n^i\right)$ converges to x^i in $(\mathbb{R}, d_2)$ for any $i \in \{1, 2, \ldots, k\}$.*

Here we use superscript to index coordinates of vectors, since we have used subscript to index terms of the sequences.

Proof. $\Rightarrow$:

Take any $i \in \{1, 2, \ldots, k\}$. WTS: $x_n^i \to x^i$ in $(\mathbb{R}, d_2)$.

Take any $\varepsilon > 0$, we want to find N^i s.t. $d_2\left(x_n^i, x^i\right) < \varepsilon$ for any $n > N^i$.

Because $x_n \to x$, there exists N s.t. $d_2(x_n, x) < \varepsilon$ for any $n > N$. Let $N^i := N$, and I claim that this is an N^i we need to find. This is because for any $n > N^i := N$, we have

$$\begin{aligned} d_2\left(x_n^i, x^i\right) &= \left|x_n^i - x^i\right| = \sqrt{\left(x_n^i - x^i\right)^2} \\ &\leq \sqrt{\sum_{j=1}^{k}\left(x_n^j - x^j\right)^2} = d_2(x_n, x) < \varepsilon \end{aligned}$$

$\Leftarrow$:

Take any $\varepsilon > 0$, we want to find N s.t. $d_2(x_n, x) < \varepsilon$ for any $n > N$.

Because $x_n^i \to x^i$, there exists N^i s.t. $d_2\left(x_n^i, x^i\right) < \varepsilon/\sqrt{k}$ for any $n > N^i$. Let $N := \max\{N_1, \ldots, N_k\}$, and I claim that this is an N we want to find. This is because for any $n > N$, we have $n > N^i$ and thus $d_2\left(x_n^i, x^i\right) < \varepsilon/\sqrt{k}$ for any i, and therefore

$$d_2(x_n, x) = \sqrt{\sum_{j=1}^{k}\left(x_n^j - x^j\right)^2} < \sqrt{k\left(\varepsilon/\sqrt{k}\right)^2} = \varepsilon$$

$\square$

The following proposition and its corollary state that the operators +, −, ×, and / preserves limit in $\mathbb{R}$.

Proposition 2.7. *In* $(\mathbb{R}, d_2)$*, let there be two convergent sequences* $x_n \to x$ *and* $y_n \to y$*. Then*
(1) $x_n + y_n \to x + y$*,*
(2) $x_n y_n \to xy$*, and*
(3) if $x \neq 0$*, then* $1/x_n \to 1/x$

Proof. Let's prove (2), and leave (1) and (3) as exercises.

Take any $\varepsilon > 0$, I want to find N s.t. $|x_n y_n - xy| < \varepsilon$ for any $n > N$.

Because (y_n) is convergent, it is bounded, i.e. there exists an open ball $(z - r, z + r)$ that contains $\{y_1, y_2, \ldots\}$. Let $M := \max\big\{|z - r|, |z + r|\big\}$, and by construction $|y_n| < M$ for any n.

Because $x_n \to x$, there exists N_x s.t. $|x_n - x| < \varepsilon/2M$. Because $y_n \to y$, there exists N_y s.t. $|y_n - y| < \varepsilon/2\big(|x| + 1\big)$.

Let $N := \max\big\{N_x, N_y\big\}$, and I claim that this is an N we need to find. This is because for any $n > N$, we have

$$\begin{aligned}
|x_n y_n - xy| &= \big|(x_n - x)\, y_n + (y_n - y)\, x\big| \\
&\leq |x_n - x| \cdot |y_n| + |y_n - y| \cdot |x| \\
&< \frac{\varepsilon}{2M} \cdot M + \frac{\varepsilon}{2\big(|x| + 1\big)} \cdot |x| \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\end{aligned}$$

$\square$

Corollary 2.8. *In* $(\mathbb{R}, d_2)$*, consider two convergent sequences* $x_n \to x$ *and* $y_n \to y$*. Then*

1. $x_n - y_n \to x - y$*,*
2. *if* $y \neq 0$*, then* $x_n/y_n \to x/y$

Proof.
1. Clearly, the constant sequence $z_n := -1$ converges to $z = -1$. Therefore, by Proposition 2.7(2), we have $z_n y_n \to zy$. So

$$-y_n = z_n y_n \to zy = -y$$

By Proposition 2.7(1), we have

$$x_n - y_n = x_n + (-y_n) \to x + (-y) = x - y$$

2. By Proposition 2.7(3), we have $1/y_n \to 1/y$. Then by Proposition 2.7(2), we have

$$x_n/y_n = x_n \cdot \big(1/y_n\big) \to x \cdot \big(1/y\big) = x/y$$

$\square$

When we combine Proposition 2.7 with Proposition 2.6, we can obtain similar results for convergence of vectors. For example, if $x_n \to x$ and $y_n \to y$ in $\left(\mathbb{R}^k, d_2\right)$, then we have $x_n + y_n \to x + y$ in $\left(\mathbb{R}^k, d_2\right)$.

2.3 Monotone Convergence and Bolzano-Weierstrass Theorem

Because sequences are special cases of functions, the definition of monotonicity of functions applies to sequences. A sequence is **increasing** iff $x_m \le x_n$ for any $m \le n$; is **decreasing** iff $x_m \ge x_n$ for any $m \le n$; is **monotone** iff it is increasing or decreasing. We can also define the strict versions of increasing/decreasing sequences in the natural way as we did for functions.

Now let's use the least upper bound property of $\mathbb{R}$ to prove the following important theorem.

Theorem 2.9 (Monotone Convergence Theorem)**.** *Every increasing (resp. decreasing) and bounded from above (resp. below) real sequence* (x_n) *is convergent in* $(\mathbb{R}, d_2)$*.*

Notice that an increasing/decreasing sequence is automatically bounded from below/above by its first term. Therefore, the theorem can also be stated that every monotone and bounded real sequence is convergent in $(\mathbb{R}, d_2)$.

Proof. (1) Take any increasing and bounded from above real sequence (x_n). Because the range of the sequence $\{x_1, x_2, \ldots\}$ is bounded from above, by l.u.b. property of $\mathbb{R}$, it has a least upper bound.

Let $x := \sup\{x_1, x_2, \ldots\}$, and we WTS $x_n \to x$.

Take any $\varepsilon > 0$. We want to find N s.t. $|x_n - x| < \varepsilon$ for any $n > N$. Because x is the least upper bound of $\{x_1, x_2, \ldots\}$, $x - \varepsilon$ is not an upper bound, and therefore there exists N s.t. $x_N > x - \varepsilon$. Therefore, for any $n > N$, we have

$$x \ge x_n \ge x_N > x - \varepsilon$$

and therefore $|x_n - x| < \varepsilon$.

(2) Take any decreasing and bounded from below real sequence (x_n). Clearly $(-x_n)$ is increasing and bounded from above. By (1) we have $(-x_n)$ is convergent, and thus (x_n) is also convergent. $\square$

Given a sequence (x_n), a **subsequence** of (x_n) is a sequence $\left(x_{n_k}\right)$ indexed by $k \in \mathbb{N}$, where (n_k) is a strictly increasing sequence in $\mathbb{N}$. For example, if the sequence (n_k) is $2, 4, 5, 9, \ldots$, then the subsequence $\left(x_{n_k}\right)$ is $x_2, x_4, x_5, x_9, \ldots$.

Lemma 2.10. *Every sequence in* $\mathbb{R}$ *has a monotone subsequence.*

Proof. Take any sequence (x_n) in $\mathbb{R}$. Call the term x_n a dominant term if $x_n \ge x_m$ for any $m \ge n$.

Case 1: (x_n) has infinitely many dominant terms

Then these dominant terms constitute a decreasing subsequence.

Case 2: (x_n) has finitely many dominant terms

Let x_N be the last dominant term. Let $n_1 = N + 1$, and so x_{n_1} is not a dominant term. By definition, there exists $n_2 > n_1$ s.t. $x_{n_2} > x_{n_1}$. The term x_{n_2} itself is not dominant either, and so there exists $n_3 > n_2$ s.t. $x_{n_3} > x_{n_2}$... Therefore, we obtain a strictly increasing subsequence.

Case 3: (x_n) has no dominant term

Let $n_1 = 1$, and construct a strictly increasing subsequence as in Case 2.

Therefore, we can always find a monotone subsequence. $\square$

If we combine the lemma we have just proved with Monotone Convergence Theorem, we immediately obtain the **Bolzano-Weierstrass theorem**: every bounded sequence in $(\mathbb{R}, d_2)$ has a convergent subsequence.

We can easily extend Bolzano-Weierstrass theorem to $\left(\mathbb{R}^k, d_2\right)$. So we have proved the following theorem.

Theorem 2.11 (Bolzano-Weierstrass)**.** *Every bounded sequence in* $\left(\mathbb{R}^k, d_2\right)$ *has a convergent subsequence.*

Proof. (sketch)

Take any bounded vector sequence in $\left(\mathbb{R}^k, d_2\right)$. Clearly, each of its coordinate is a bounded real sequence in $(\mathbb{R}, d_2)$. Then we can apply Bolzano-Weierstrass theorem in $(\mathbb{R}, d_2)$ to find a convergent subsequence for the first coordinate. Then we can find a subsequence of this subsequence that is convergent in the second coordinate. Repeat this process, and we finally obtain a subsequence that is convergent in every coordinate. By Proposition 2.6, the subsequence for the vector is convergent. $\square$

3 Compactness

Compactness is a stronger notion than closedness. It is also a crucial concept, because compact sets have many desirable properties that closed sets don't have.

Definition 3.1. *Let* (X, d) *be a metric space, and* S *a subset of* X*. A family of open sets* $\{E_\alpha\}_{\alpha \in A}$ *is an* ***open cover*** *of* S *iff* $\bigcup_{\alpha \in A} E_\alpha \supset S$.

Definition 3.2. *Let* (X, d) *be a metric space, and* S *a subset of* X*. The set* S *is* ***compact*** *iff* $\forall$ *open cover* $\{E_\alpha\}_{\alpha \in A}$ *of* S*,* $\exists$ *a finite* $B \subset A$ *s.t.* $\{E_\alpha\}_{\alpha \in B}$ *is also an open cover of* S*.*

To illustrate the definition of compact sets, let's verify that the open interval $(0, 1)$ is not a compact set in $(\mathbb{R}, d_2)$. To do this, it is sufficient to provide an open cover that does not have a finite subcover. Consider the family of open sets $\left\{\left(1/n, 1 - 1/n\right)\right\}_{n=3}^{+\infty}$. This covers $(0, 1)$ because any point strictly between 0 and 1 will be eventually covered by $\left(1/n, 1 - 1/n\right)$ when n is large enough. There is no finite subcover, since any finite family of $\left(1/n, 1 - 1/n\right)$ has a largest one, and it does not cover $(0, 1)$.

By definition, the concept of compactness relies on the metric space we are working with, just like openness and closedness. A set can be compact in one metric space, but not in another metric space. However, compactness behaves much better than openness and closedness, in the sense that enlarging or shrinking the whole space does not affect compactness as long as we use the same metric. This result is formulated below.

Proposition 3.3. *Let* (X, d) *be a metric space, and* $S \subset Y \subset X$*. Then* S *is compact in* (X, d) *iff* S *is compact in* (Y, d)*.*

Recall that (Y, d) in fact means $\left(Y, d|_Y\right)$, rigorously speaking.

See Theorem 2.33 in Rudin for a proof.

In the proposition above, if we let $Y := S$, we have that S is compact in (X, d) iff S is compact in (S, d). A metric space (X, d) is said to be a **compact metric space** iff X is a compact set in (X, d). So we know that if S is compact in (X, d), then (S, d) itself is a compact metric space, and vice versa.

3.1 General Properties

The theorem below states that compactness is stronger than closedness.

Theorem 3.4. *Let* (X, d) *be a metric space, and* S *a subset of* X*. If* S *is compact in* (X, d)*, then* S *is closed in* (X, d)*.*

Proof. WTS: S^c is open in (X, d)

Take any $x \in S^c$, we want to find $r > 0$ s.t. $B_r(x) \subset S^c$.

Take any $y \in S$, let $r_y := d(y, x)/2$. Then clearly $B_{r_y}(x)$ and $B_{r_y}(y)$ are disjoint.

Notice that $\left\{B_{r_y}(y)\right\}_{y \in S}$ is an open cover of S. By compactness of S, there exists $\{y_1, y_2, \dots, y_n\}$ s.t. $\left\{B_{r_{y_i}}(y_i)\right\}_{i=1}^{n}$ is also an open cover of S.

Let $r := \min\left\{r_{y_1}, r_{y_2}, \dots, r_{y_n}\right\}$. WTS: $B_r(x) \subset S^c$.

Clearly $B_r(x)$ is disjoint with $B_{r_{y_i}}(y_i)$ for any i. So $B_r(x)$ is disjoint with the union of $B_{r_{y_i}}(y_i)$'s, and thus $B_r(x)$ is disjoint with S, which implies $B_r(x) \subset S^c$. $\square$

The next theorem states that a compact set must be bounded.

Theorem 3.5. *Let (X, d) be a metric space, and S a subset of X. If S is compact in (X, d), then S is bounded in (X, d).*

Proof. Arbitrarily take a point $x \in X$. Clearly, $\left\{B_n(x)\right\}_{n=1}^{\infty}$ is an open cover of S, because any $y \in X$ has a finite distance to x, and will be eventually covered by $B_n(x)$ when n is large enough. By compactness of S, there exists $\{n_1, n_2, \dots, n_k\}$ s.t. $\left\{B_{n_i}(x)\right\}_{i=1}^{k}$ is also an open cover of S.

Let $r := \max\{n_1, n_2, \dots, n_k\}$. Then $B_r(x) \supset S$, and so S is bounded. $\square$

Combining the Theorem 3.4 and 3.5, we conclude that a compact set must be closed and bounded.

The next theorem provides a way to prove compactness. It states that a closed set contained in a compact set is also compact. Therefore, in order to show that S is compact, we can instead show that S is closed and that some other set containing S is compact.

Theorem 3.6. *Let (X, d) be a metric space, and $S \subset Y \subset X$. If S is closed in (X, d) and Y is compact in (X, d), then S is compact in (X, d).*

Proof. Take any open cover $\{E_\alpha\}_{\alpha \in A}$ of S. We want to find a finite family chosen from $\{E_\alpha\}_{\alpha \in A}$ that also covers S.

Clearly, $\{E_\alpha\}_{\alpha \in A} \cup \{S^c\}$ covers the whole space, and thus covers Y. Because Y is compact, there exists a finite family chosen from $\{E_\alpha\}_{\alpha \in A} \cup \{S^c\}$ that covers Y. Because $S \subset Y$, the finite family also covers S. If the finite family contains S^c, then we can remove it from the family, then the family still covers S, since S^c has no contribution to covering S. So we have obtained a finite family chosen from $\{E_\alpha\}_{\alpha \in A}$ that covers S. $\square$

Theorem 3.4 and 3.6 together implies that in compact metric spaces, closedness and compactness are equivalent.

The discussions above apply to general metric spaces. The next subsection is dedicated to Euclidean spaces $\left(\mathbb{R}^k, d_2\right)$, and establishes more results.

3.2 Heine-Borel Theorem in $\left(\mathbb{R}^k, d_2\right)$

In general metric spaces, we have shown that a compact set must be closed and bounded. In Euclidean spaces $\left(\mathbb{R}^k, d_2\right)$, the reverse is also true, i.e. a closed and bounded set in $\left(\mathbb{R}^k, d_2\right)$ must be compact. Therefore, in Euclidean spaces, compactness is equivalent to closedness plus boundedness, and this equivalence is known as Heine-Borel theorem.

We establish this result in several steps.

Lemma 3.7. *Any closed interval $[a, b]$ is compact in $(\mathbb{R}, d_2)$.*

Proof. Take any closed interval $[a, b]$, and suppose that it is not compact. Then there exists an open cover $\{E_\alpha\}_{\alpha\in A}$ of $[a, b]$ without a finite subcover.

Let $a_0 := a$, and $b_0 := b$.

Cut the interval $[a_0, b_0]$ in half: $\left[a_0, (a_0 + b_0)/2\right]$ and $\left[(a_0 + b_0)/2, b_0\right]$. At least one of them cannot be finitely covered (otherwise the interval $[a_0, b_0]$ can be finitely covered). Take the one that cannot be finitely covered, and label it as $[a_1, b_1]$.

Cut the interval $[a_1, b_1]$ in half: $\left[a_1, (a_1 + b_1)/2\right]$ and $\left[(a_1 + b_1)/2, b_1\right]$. At least one of them cannot be finitely covered. Take the one that cannot be finitely covered, and label it as $[a_2, b_2]$.

Repeat this process, and we get a shrinking sequence of intervals $[a_0, b_0] \supset [a_1, b_1] \supset \cdots$, and each of them cannot be finitely covered using the open cover $\{E_\alpha\}_{\alpha\in A}$.

Because (a_n) is increasing and bounded from above by b_0, (a_n) converges to some limit a^*. Symmetrically, (b_n) converges to some limit b^*. Because $b_n - a_n = \left(1/2\right)^n (b - a) \to 0$, we know that

$$b_n = a_n + (b_n - a_n) \to a^* + 0 = a^*$$

and therefore $b^* = a^*$. That is, the sequence of intervals $[a_0, b_0] \supset [a_1, b_1] \supset \cdots$ shrinks to one point a^*. Because $a^* \in [a, b]$, it is covered by some open set E_{α^*} in the open cover.

Therefore, there exists $B_r(a^*) \subset E_{\alpha^*}$. Because (a_n) and (b_n) both converge to a^*, there exists $\hat{n}$ s.t. $a_{\hat{n}}, b_{\hat{n}} \in B_r(a^*)$, and therefore $[a_{\hat{n}}, b_{\hat{n}}] \subset B_r(a^*) \subset E_{\alpha^*}$. So $[a_{\hat{n}}, b_{\hat{n}}]$ can be finitely covered using the open cover $\{E_\alpha\}_{\alpha\in A}$, which contradicts the construction of the sequence $\left([a_n, b_n]\right)$. □

It is not difficult to extend the lemma to $\left(\mathbb{R}^k, d_2\right)$.

Lemma 3.8. *Every k-cell $[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_k, b_k]$ is compact in $\left(\mathbb{R}^k, d_2\right)$.*

We can use the same idea to prove this result for $\mathbb{R}^k$ as that for $\mathbb{R}$. Take any k-cell, and suppose that it is not compact. Then there exists an open cover $\{E_\alpha\}_{\alpha\in A}$ of the cell without a finite subcover. Then in each coordinate of the k-cell, we cut the interval in half, and in total we get 2^k sub-cells. At least one of them cannot be finitely covered. Take the one that cannot be finitely covered, and cut it into 2^k sub-cells, and repeat this process. We can get a shrinking sequence of k-cells, each of whom cannot be finitely covered. It can be shown that the sequence shrinks to one limit point x^* in the original k-cell. Then let the open set in $\{E_\alpha\}_{\alpha\in A}$ that covers x^* be E_{α^*}. It can be shown that the sequence of k-cells will eventually go into E_{α^*}, and thus can be finitely covered. This contradicts the construction of this sequence. See Theorem 2.40 in Rudin for details.

Now let's consider a closed and bounded set S in $\left(\mathbb{R}^k, d_2\right)$. By definition of boundedness, S can be bounded by an open ball. Clearly, an open ball in $\left(\mathbb{R}^k, d_2\right)$ can be bounded by a k-cell. So S is a subset of a k-cell, which is compact in $\left(\mathbb{R}^k, d_2\right)$ by the lemma above. Furthermore, because S is assumed to be closed in $\left(\mathbb{R}^k, d_2\right)$, by Theorem 3.6, we know that S is compact itself. Therefore we have proved the following theorem.

Theorem 3.9. *In $\left(\mathbb{R}^k, d_2\right)$, a closed and bounded set S must be compact.*

Combining this theorem with Theorem 3.4 and 3.5, we have the well-known Heine-Borel theorem.

Theorem 3.10 (Heine-Borel)**.** *In $\left(\mathbb{R}^k, d_2\right)$, a set S is compact iff it is closed and bounded.*

Heine-Borel theorem states that in $\left(\mathbb{R}^k, d_2\right)$, to check whether a set is compact, we can instead check whether the set is closed and bounded. This greatly simplifies our job, since the definition of compactness involving open covers is not easy to check in most cases.

Keep in mind that Heine-Borel theorem only works in Euclidean spaces $\left(\mathbb{R}^k, d_2\right)$. In general metric spaces (X, d), although compactness always implies closedness plus boundedness, the reverse is not true in general. For example, in $(\mathbb{R}_{++}, d_2)$, the set $(0, 1]$ is closed and bounded, but not compact (consider the open cover $\left\{(1/n, +\infty)\right\}_{n=1}^{\infty}$). For another example, in $(\mathbb{R}, d)$, where d is the discrete metric, the set $[0, 1]$ is closed and bounded, but not compact. In fact, under the discrete metric, a set is compact iff it is finite (exercise).

3.3 Sequential Compactness

There is another notion of compactness, called sequential compactness.

Definition 3.11. *Let (X, d) be a metric space, and S a subset of X. The set S is* ***sequentially compact*** *iff any sequence (x_n) in S has a subsequence convergent to some $x^* \in S$.*

Theorem 3.12. *Let (X, d) be a metric space, and S a subset of X. The set S is compact iff it is sequentially compact.*

This equivalence holds in general metric spaces[3], but here we will just provide the proof in Euclidean spaces $\left(\mathbb{R}^k, d_2\right)$.

Proof. "$\Rightarrow$":

If S in $\left(\mathbb{R}^k, d_2\right)$ is compact, then it is bounded. Then any sequence (x_n) in S must be bounded. By Bolzano-Weierstrass theorem, (x_n) has a subsequence convergent to some $x^* \in \mathbb{R}^k$. Then by the sequential definition of closed sets, we have $x^* \in S$. Therefore, S is sequentially compact.

"$\Leftarrow$" :

If S in $\left(\mathbb{R}^k, d_2\right)$ is sequentially compact, then it must be closed; otherwise we can find a sequence (x_n) in S convergent to some x^* outside S, and any subsequence of (x_n) must also converge to $x^* \notin S$, so it does not have a subsequence convergent to some point in S.

Also, the set S must also be bounded. Otherwise we can construct a sequence (x_n) s.t. $d_2(x_n, 0) > n$, and so (x_n) does not even have a convergent subsequence. Therefore, S is both closed and bounded, and therefore compact. $\square$

4 Cauchy Sequences and Completeness

Definition 4.1. *In metric space (X, d), a sequence (x_n) is a* ***Cauchy sequence*** *iff $\forall\ \varepsilon > 0$, $\exists\ N \in \mathbb{N}$ s.t.*

$$d(x_m, x_n) < \varepsilon$$

for any $m, n > N$.

Clearly, Cauchy sequences must be bounded just like convergent sequences. This is left as an exercise.

[3] However, the two notions are not the same in *topological spaces*, which is a generalization of metric spaces and is out of the scope of our math camp.

Proposition 4.2. *In metric space (X, d), a convergent sequence (x_n) is a Cauchy sequence.*

Proof. Let the limit of (x_n) be x.

Take any $\varepsilon > 0$. I want to find N s.t. $d(x_m, x_n) < \varepsilon$ for any $m, n > N$.

Because $x_n \to x$, there exists N s.t. $d(x_k, x) < \varepsilon/2$ for any $k > N$. Therefore, for any $m, n > N$, we have

$$d(x_m, x_n) \leq d(x_m, x) + d(x_n, x) < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

□

However, a Cauchy sequence may fail to be convergent, for example the sequence $\left(1/n\right)$ in $(\mathbb{R}_{++}, d_2)$. If the metric space (X, d) has the property that every Cauchy sequence converges, then we call it a complete metric space. The metric space $(\mathbb{R}_{++}, d_2)$ is not complete.

Definition 4.3. *Let (X, d) be a metric space, and S a subset of X. The set S is a **complete** set iff any Cauchy sequence in S converges to a limit point in S.*

*A metric space (X, d) is a **complete metric space** iff X is a complete set in (X, d).*

Completeness is stronger than closedness, but weaker than compactness.

A complete set S must be closed. Otherwise we can find a sequence (x_n) in S convergent to a point outside S. Because (x_n) is Cauchy, this contradicts the completeness of S.

The next result states that a compact set S must be complete.

Proposition 4.4. *Let (X, d) be a metric space, and S a subset of X. If the set S is compact, then it is complete.*

Proof. Take any Cauchy sequence (x_n) in S. We want to find an $x \in S$ s.t. $x_n \to x$.

Because S is compact, and so is sequentially compact, we can find a subsequence $\left(x_{n_k}\right)$ convergent to some $x \in S$.

Now we only need to show $x_n \to x$

Take any $\varepsilon > 0$. We want to find N s.t. $d(x_n, x) < \varepsilon$ for any $n > N$.

Because (x_n) is Cauchy, there exists N s.t. $d(x_m, x_n) < \varepsilon/2$ for any $m, n > N$.

Because $x_{n_k} \to x$, there exists K s.t. $n_K > N$ and $d\left(x_{n_K}, x\right) < \varepsilon/2$.

Then for any $n > N$, we have

$$d(x_n, x) \leq d\left(x_n, x_{n_K}\right) + d\left(x_{n_K}, x\right) < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

□

In the proof above, the candidate x for Cauchy sequence limit is provided by the sequential compactness of S. Then the trick to prove $x_n \to x$ is to bind the whole sequence to the convergent subsequence using Cauchy, and then bind the subsequence to the limit using its convergence.

Using the same trick, we can show that the Euclidean spaces $\left(\mathbb{R}^k, d_2\right)$ are complete.

Proposition 4.5. *The Euclidean space $\left(\mathbb{R}^k, d_2\right)$ is a complete metric space.*

The proof of this result is left as an exercise. Notice that the candidate x for Cauchy sequence limit is provided by Bolzano-Weierstrass theorem.

4.1 Contraction Mapping Theorem

This subsection discusses an important fixed-point result in complete metric spaces, known as Contraction Mapping theorem or Banach Fixed Point theorem. This result has important implications in dynamic programming.

A function is called a **self-map** iff it maps its domain to itself, i.e. $f : X \to X$. Note that a self-map need not be surjective or injective.

A point $x^* \in X$ is called a **fixed point** of the self-map $f : X \to X$, iff $f(x^*) = x^*$. Intuitively, the fixed point x^* does not "move away" if we apply f to it.

Definition 4.6. *Let (X, d) be a metric space. A self-map $f : X \to X$ is said to be a **contraction** iff $\exists$ real $\lambda < 1$ s.t.*

$$d\Big(f(x), f(x')\Big) \leq \lambda \cdot d(x, x')$$

for any $x, x' \in X$.

Now we state the theorem.

Theorem 4.7 (Contraction Mapping Theorem)**.** *Let (X, d) be a complete metric space, and $f : X \to X$ a contraction. Then f has a unique fixed point x^*. Further, for any $x \in X$, we have $\lim_{n\to\infty} f^n(x) = x^*$.*

The notation $f^n(x)$ means to apply f to x n times, i.e. $f^2(x) := f\big(f(x)\big)$, $f^3(x) = f\Big(f\big(f(x)\big)\Big)$, and so on.

Outline of the proof:

1. Show that the sequence $\big(f^n(x_0)\big)$ is Cauchy for an arbitrary $x_0 \in X$. So by completeness of (X, d), $\big(f^n(x_0)\big)$ converges to some $x^* \in X$.
2. Show that $x^* := \lim_{n\to\infty} f^{(n)}(x_0)$ is indeed a fixed point of f.
3. Show that x^* is the unique fixed point of f.
4. Show that $f^n(x) \to x^*$ for any starting point $x \in X$.

Proof. Arbitrarily take $x_0 \in X$. Define $x_n := f^n(x_0)$, for any $n \in \mathbb{N}$.

Step 1: WTS (x_n) is Cauchy

Take any $\varepsilon > 0$, we want to find N s.t. $d(x_m, x_n) < \varepsilon$ for any $m, n > N$.

Because f is a contraction, we have $d\Big(f(x), f(x')\Big) \leq \lambda \cdot d(x, x')$ for some $\lambda < 1$.

Let N be s.t. $\lambda^N < (1-\lambda)\varepsilon / d(x_0, x_1)$. This is possible because $\lambda^N \to 0$, and $(1-\lambda)\varepsilon/d(x_0, x_1) > 0$ [4].

WTS: $d(x_m, x_n) < \varepsilon$ for any $m, n > N$.

Since

$$d(x_k, x_{k+1}) = d\big(f(x_{k-1}), f(x_k)\big) \leq \lambda d(x_{k-1}, x_k)$$

we have,

[4] If $d(x_0, x_1) = 0$, it is obvious that $f : x \to x$ is not a contraction.

$$
\begin{aligned}
d\left(x_k, x_{k+1}\right) &\leq \lambda d\left(x_{k-1}, x_k\right) \\
&\leq \lambda^2 d\left(x_{k-2}, x_{k-1}\right) \\
&\leq \cdots \\
&\leq \lambda^k d\left(x_0, x_1\right)
\end{aligned}
$$

Take any $m, n > N$. Without loss of generality, assume that $m \leq n$, and we have

$$
\begin{aligned}
d\left(x_m, x_n\right) &\leq d\left(x_m, x_{m+1}\right) + d\left(x_{m+1}, x_{m+2}\right) + \cdots + d\left(x_{n-1}, x_n\right) \\
&\leq \lambda^m d\left(x_0, x_1\right) + \lambda^{m+1} d\left(x_0, x_1\right) + \cdots + \lambda^{n-1} d\left(x_0, x_1\right) \\
&= \lambda^m d\left(x_0, x_1\right)\left(1 + \lambda + \cdots + \lambda^{n-m-1}\right) \\
&= \lambda^m d\left(x_0, x_1\right) \frac{1 - \lambda^{n-m}}{1 - \lambda} < \lambda^N d\left(x_0, x_1\right) \frac{1}{1 - \lambda} \\
&< \frac{(1 - \lambda)\varepsilon}{d\left(x_0, x_1\right)} \cdot d\left(x_0, x_1\right) \frac{1}{1 - \lambda} = \varepsilon
\end{aligned}
$$

As a result, we have shown that (x_n) is Cauchy.
Because (X, d) is a complete metric space, (x_n) converges to some limit $x^* \in X$.
Step 2: WTS $f(x^*) = x^*$
I want to show this by showing $d\left(f(x^*), x^*\right) = 0$.
It is sufficient to show that $d\left(f(x^*), x^*\right) < \varepsilon$ for any $\varepsilon > 0$.
Take any $\varepsilon > 0$.
Because $x_n \to x^*$, there exists N s.t. $d(x_n, x^*) < \varepsilon/2$ for any $n \geq N$.
Then we have

$$
\begin{aligned}
d\left(f\left(x^*\right), x^*\right) &\leq d\left(f\left(x^*\right), x_{N+1}\right) + d\left(x_{N+1}, x^*\right) \\
&= d\left(f\left(x^*\right), f\left(x_N\right)\right) + d\left(x_{N+1}, x^*\right) \\
&\leq \lambda d\left(x^*, x_N\right) + d\left(x_{N+1}, x^*\right) \\
&< \lambda \cdot \varepsilon/2 + \varepsilon/2 < \varepsilon
\end{aligned}
$$

Step 3: WTS x^* is the unique fixed point of f

Assume by contradition that $\exists \hat{x} \in S, \hat{x} \neq x^*$ such that $f(\hat{x}) = \hat{x}$. Then $d\left(f(\hat{x}), f(x^*)\right) = d(\hat{x}, x^*) > 0$, i.e. $\lambda = 1$. Therefore f is not a contraction. Contradiction!

Step 4: WTS $f^n(x) \to x^*$ for any starting point $x \in X$

Take $\forall y_0 \in S, y_0 \neq x_0$ and $y^* := \lim_{n\to\infty} f^{(n)}(y_0) \in S$ (from Step 1 we know that such y^* exists). WTS $y^* = x^*$, i.e. $d(x^*, y^*) = 0$. It suffices to show that $d(y^*, x^*) < \varepsilon$ for any $\varepsilon > 0$.

Denote $y_n := f^{(n)}(y_0)$. Let N_1 be such that $\lambda^{N_1} < \varepsilon/(3d(x_0, y_0))$. Then we have $d(x_n, y_n) \leq \lambda^n d(x_0, y_0) \leq \lambda^{N_1} d(x_0, y_0) < \varepsilon/3$, $\forall n \geq N_1$.

Also, as we have $x_n \to x^*, y_n \to y^*$, there exists N_2 s.t. $d(x_n, x^*) < \varepsilon/3$, $d(y_n, y^*) < \varepsilon/3$ for any $n \geq N_2$.

Take $N := \max\{N_1, N_2\}$, then for any $n \geq N$, we have:

$$
\begin{aligned}
d(x^*, y^*) &\leq d(x_n, x^*) + d(y_n, y^*) + d(x_n, y_n) \\
&< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon
\end{aligned}
$$

OR: Step 4 directly follows from Steps 1,2 and 3, because by Step 1 for any $y \in X$ we have $y^* := \lim_{n\to\infty} f^{(n)}(y)$ exists and by Step 2 $y^* = f(y^*)$. By Step 3 we see that $y^* = a \in X$ for $\forall y \in X$. $\square$

The completeness of the metric space plays a central role in Contraction Mapping theorem, because it gives us the limit x^* of the sequence $\left(f^n(x)\right)$, which turns out to be the fixed point of f we are searching for.

Without completeness of the metric space, the result is not true. For example, in $\left(\mathbb{R}\backslash\{0\}, d_2\right)$, the function $f(x) = x/2$ is a contraction, but it does not have a fixed point.

5 Continuity of Functions

5.1 Limits of Functions

So far we defined the notion of limit for sequences. Now let's do it for functions.

Definition 5.1. *Let (X, d_X) and (Y, d_Y) be metric spaces. Let S be a subset of X, function $f : S \to Y$, and $x_0 \in S'$ (a limit point of S).*

We say that y_0 is a ***limit of f at*** *x_0, iff $\forall\ \varepsilon > 0$, $\exists\ \delta > 0$ s.t.*

$$f((B_\delta(x_0) \cap S)\backslash\{x_0\}) \subset B_\varepsilon(y_0).$$

In this case, we denote $\lim_{x\to x_0} f(x) = y_0$.

Notice that x_0 has to be a limit point of the domain S, since we have to make sure that there exists $x \in S$ s.t. $0 < d(x, x_0) < \delta$, no matter how small δ is. However, x_0 is allowed to be outside the domain S[5]. For example, consider the function $f : (0, 1) \to \mathbb{R}$ defined as $f(x) = 2x$. It makes sense to talk about the limit of f at 1, although 1 is not in the domain of f. In this case, $\lim_{x\to1} f(x) = 2$. (In $\mathbb{R}^k$, we always use the Euclidean distance d_2 by default, unless stated otherwise.)

The concept of the limit of f at x_0 has nothing to do with the value of f at x_0. Instead, it captures the behavior of the function f only nearby x_0 but not at x_0. For example, consider the function $f : \mathbb{R} \to \mathbb{R}$ defined as

$$f(x) := \begin{cases} 2x, \text{ if } x \neq 1 \\ 0, \text{ if } x = 1 \end{cases}$$

Notice that $f(1) = 0$, but $\lim_{x\to1} f(x) = 2$. (Again, d_2 is used by default.)

Similar to the limit of a sequence, we can use triangle inequality to show that limit of f at x_0 is unique, if exists. This enables us to talk about "the" limit of f at x_0, and use the notation $\lim_{x\to x_0} f(x)$ without ambiguity.

The next theorem reveals the relation of the limit of a function to the limit of sequences.

Theorem 5.2. *Let (X, d_X) and (Y, d_Y) be metric spaces. Let S be a subset of X, function $f : S \to Y$, and $x_0 \in S'$. Then the limit of f at x_0 is y_0 iff the sequence $\left(f(x_n)\right)$ converges to y_0 for any sequence (x_n) in $S\backslash\{x_0\}$ that converges to x_0.*

[5] To allow x_0 to be outside the domain S is an important generality to maintain. We know that the derivative of a single variable function f at x_0 is defined as

$$f'(x_0) := \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

Clearly, the point x_0 is not in the domain of the slope function $s(x) := \frac{f(x)-f(x_0)}{x-x_0}$, since the denominator cannot be 0. However, we are still able to talk about the limit of $s(x)$ at x_0.

Proof. $\Rightarrow$:

Take any sequence (x_n) in S convergent to x_0 s.t. $x_n \neq x_0$ for any n.

WTS: $f(x_n) \to y_0$

Take any $\varepsilon > 0$. We want to find N s.t. $d_Y\big(f(x_n), y_0\big) < \varepsilon$ for any $n > N$.

Because $\lim_{x\to x_0} f(x) = y_0$, there exists $\delta > 0$ s.t. $d_Y\big(f(x), y_0\big) < \varepsilon$ for any $x \in S$ with $0 < d_X(x, x_0) < \delta$.

Because $x_n \to x_0$, there exists N s.t. $d_X(x_n, x_0) < \delta$ for any $n > N$.

Take any $n > N$, because $x_n \neq x_0$, we have $0 < d_X(x_n, x_0) < \delta$, and so $d_Y\big(f(x_n), y_0\big) < \varepsilon$.

$\Leftarrow$:

Suppose that y_0 is not a limit of f at x_0. Then there exists $\hat{\varepsilon} > 0$ s.t. there is no $\delta > 0$ s.t. $d_Y\big(f(x), y_0\big) < \hat{\varepsilon}$ for any $x \in S$ with $0 < d_X(x, x_0) < \delta$.

Then for any $n \in \mathbb{N}$, we can find $x_n \in S$ s.t. $0 < d_X(x_n, x_0) < 1/n$, but $d_Y\big(f(x_n), y_0\big) \geq \hat{\varepsilon}$.

Clearly, the sequence $x_n \to x_0$, and $x_n \neq x_0$ for any n, but $\big(f(x_n)\big)$ does not converge to y_0. This contradicts our assumption. $\square$

Because of the close link of the function limit to sequence limit, as revealed by the theorem above, Proposition 2.6 and Proposition 2.7 also works for function limit. We state them as the following two propositions.

Proposition 5.3. *Let (X, d_X) be a metric space, S be a subset of X, and $x_0 \in S'$. Let f be a function from S to $\mathbb{R}^k$. Then $\lim_{x\to x_0} f(x)$ exists iff $\lim_{x\to x_0} f_i(x)$ exists for any $i = 1, 2, \dots k$. Furthermore, when the limit exist, the i-th coordinate of $\lim_{x\to x_0} f(x)$ is equal to $\lim_{x\to x_0} f_i(x)$.*

Note that f_i is used to denote the i-th coordinate of f, and so f_i is a function from S to $\mathbb{R}$.

Proposition 5.4. *Let (X, d_X) be a metric space, S be a subset of X, and $x_0 \in S'$. Let f and g be functions from S to $\mathbb{R}$ s.t. $\lim_{x\to x_0} f(x)$ and $\lim_{x\to x_0} g(x)$ exist. Then*

(1) $\lim_{x\to x_0} \big[f(x) + g(x)\big] = \lim_{x\to x_0} f(x) + \lim_{x\to x_0} g(x)$,

(2) $\lim_{x\to x_0} \big[f(x) g(x)\big] = \lim_{x\to x_0} f(x) \cdot \lim_{x\to x_0} g(x)$

(3) $\lim_{x\to x_0} \big(1/f\big) = 1/\lim_{x\to x_0} f(x)$, *if* $\lim_{x\to x_0} f(x) \neq 0$.

5.2 Continuity

Definition 5.5. *Let (X, d_X) and (Y, d_Y) be metric spaces. Let S be a subset of X, function $f : S \to Y$, and $x_0 \in S$.*

The function f is said to be ***continuous at*** *x_0 iff $\forall\ \varepsilon > 0$, $\exists\ \delta > 0$ s.t.*

$$f(B_\delta(x_0) \cap S) \subset B_\varepsilon(f(x_0)).$$

The function f is said to be a ***continuous function*** *iff f is continuous at x_0 for all $x_0 \in S$.*

Here we allow $x = x_0$, which is different from the definition of $\lim_{x\to x_0} f(x)$. Also notice that we require x_0 to be in the domain S of the function f (otherwise $f(x_0)$ is not defined), but not necessarily a limit point of the domain. In fact, if $x_0 \in S$ is not a limit point of S, i.e. x_0 is an isolated point of S, then f is continuous at x_0 by definition.

The relation of continuity to the limit of functions is stated below.

Proposition 5.6. *Let (X, d_X) and (Y, d_Y) be metric spaces. Let S be a subset of X, function $f : S \to Y$, and $x_0 \in S \cap S'$. Then the function f is continuous at x_0 iff $\lim_{x\to x_0} f(x) = f(x_0)$.*

Notice that this equivalence only works for x_0's that are both in the domain (to ensure "f is continuous at x_0" is defined) and are a limit point of the domain (to ensure $\lim_{x\to x_0} f(x)$ is defined).

Theorem 5.7. *Let (X, d_X) and (Y, d_Y) be metric spaces. Let S be a subset of X, function $f : S \to Y$, and $x_0 \in S$. Then the function f is continuous at x_0 iff $f(x_n) \to f(x_0)$ for any sequence (x_n) in S convergent to x_0.*

Simply put, f is continuous iff $x_n \to x_0$ implies $f(x_n) \to f(x_0)$.

Proof. $\Rightarrow$:

Take any sequence (x_n) in S convergent to x_0.

WTS: $f(x_n) \to f(x_0)$

Take any $\varepsilon > 0$. We want to find N s.t. $d_Y\big(f(x_n), f(x_0)\big) < \varepsilon$ for any $n > N$.

Because f is continuous at x_0, there exists $\delta > 0$ s.t. $d_Y\big(f(x), f(x_0)\big) < \varepsilon$ for any $x \in S$ with $d_X(x, x_0) < \delta$.

Because $x_n \to x_0$, there exists N s.t. $d_X(x_n, x_0) < \delta$ for any $n > N$.

Take any $n > N$, we have $d_X(x_n, x_0) < \delta$, and so $d_Y\big(f(x_n), f(x_0)\big) < \varepsilon$.

$\Leftarrow$:

If $x_0 \notin S'$, then f is continuous at x_0 by definition.

If $x_0 \in S'$, then by Theorem 5.2, the condition "$f(x_n) \to f(x_0)$ for any sequence (x_n) in S convergent to x_0" implies that $\lim_{x \to x_0} f(x) = f(x_0)$. Then by Proposition 5.6, f is continuous at x_0. $\square$

The theorem above reveals a direct link of continuity to convergence of sequences. Therefore, many results in convergence of sequences also apply here. For example, the following results are counterparts of Proposition 2.6 and 2.7.

Proposition 5.8. *Let (X, d_X) be a metric space, S be a subset of X, and $x_0 \in S$. Consider a function $f : S \to \mathbb{R}^k$. Then f is continuous at x_0 iff $f_i : S \to \mathbb{R}$ is continuous at x_0 for any $i = 1, 2, \ldots, k$.*

Proposition 5.9. *Let (X, d_X) be a metric space, S be a subset of X, and $x_0 \in S$. Let f and g be functions from S to $\mathbb{R}$ that are continuous at x_0. Then*

(1) $f + g$ is continuous at x_0,

(2) $f \cdot g$ is continuous at x_0, and

(3) $1/f$ is continuous at x_0 if $f(x_0) \neq 0$.

The next theorem shows that the compound of two continuous function is also continuous.

Theorem 5.10. *Let (X, d_X), (Y, d_Y), (Z, d_Z) be metric spaces. Let set S be a subset of X, and $f : S \to Y$ be a continuous function. Let T be a set s.t. $f(S) \subset T \subset Y$, and $g : T \to Z$ be a continuous function. Then $g \circ f : S \to Z$ is a continuous function.*

The proof is left as an exercise.

It can be shown that many commonly used functions, such as x^α, $\ln x$, e^x, and $\sin x$, are all continuous in their domain. Since by the two results above continuity is preserved under addition, multiplication, and compounding, then roughly speaking, all functions constructed using those "common" functions are continuous.

The next theorem provides yet another equivalent definition of continuity, which is known as the topological definition of continuous functions.

Theorem 5.11. *Let (X, d_X) and (Y, d_Y) be metric spaces, and function $f : X \to Y$. The function f is a continuous function iff $f^{-1}(E)$ is open in (X, d_X) for any set E open in (Y, d_Y).*

Simply put, f is continuous iff its inverse image of an open set is also open.

Different from the previous sequential characterization that works for functions continuous at a single particular point, this characterization only works for functions that are continuous everywhere.

Proof. $\Rightarrow$:

Take any $x \in f^{-1}(E)$. We want to find $\delta > 0$ s.t. $B_\delta(x) \subset f^{-1}(E)$.

Because $x \in f^{-1}(E)$, we have $f(x) \in E$. Because E is open in (Y, d_Y), there exists $\varepsilon > 0$ s.t. $B_\varepsilon\big(f(x)\big) \subset E$.

Because f is continuous at x, there exists $\delta > 0$ s.t. $f\big(B_\delta(x)\big) \subset B_\varepsilon\big(f(x)\big) \subset E$. Therefore, we have $B_\delta(x) \subset f^{-1}(E)$.

$\Leftarrow$:

Take any $\varepsilon > 0$. We want to find $\delta > 0$ s.t. $f\big(B_\delta(x)\big) \subset B_\varepsilon\big(f(x)\big)$.

Because $B_\varepsilon\big(f(x)\big)$ is open in (Y, d_Y), the set $f^{-1}\Big(B_\varepsilon\big(f(x)\big)\Big)$ is open in (X, d_X). Because $f(x) \in B_\varepsilon\big(f(x)\big)$, we have $x \in f^{-1}\Big(B_\varepsilon\big(f(x)\big)\Big)$. Therefore, the exists $\delta > 0$ s.t. $B_\delta(x) \subset f^{-1}\Big(B_\varepsilon\big(f(x)\big)\Big)$. As a result, $f\big(B_\delta(x)\big) \subset B_\varepsilon\big(f(x)\big)$. $\square$

As a corollary, f is continuous iff its inverse image of a closed set is also closed. To prove this, we only need to use the fact that the complement of an open set is closed, and that $f^{-1}(E^c) = \big(f^{-1}(E)\big)^c$.

Although taking the inverse image of a continuous function preserves openness and closedness, the image of an open set (or closed set) may not be open (or closed). For example, consider the function $f : \mathbb{R} \to \mathbb{R}$ defined as $f(x) := x^2$. The image of the open set $(-1, 1)$ under f is $[0, 1)$, which is not open in the codomain $\mathbb{R}$. For another example, consider the function $g : \mathbb{R}_{++} \to \mathbb{R}$ defined as $g(x) := 1/x$. The image of the closed set $[1, +\infty)$ under g is $(0, 1]$, which is not closed in the codomain $\mathbb{R}$.

5.3 Weierstrass Theorem

The following theorem states that a continuous image of a compact set is also compact.

Theorem 5.12. *Let (X, d_X) and (Y, d_Y) be metric spaces, and function $f : X \to Y$ is continuous. Then $f(K)$ is compact in (Y, d_Y) for any K compact in (X, d_X).*

Proof. Take any K compact in (X, d_X). WTS: $f(K)$ is compact in (Y, d_Y)

Take any open cover $\{E_\alpha\}_{\alpha \in A}$ of $f(K)$. We want to find a finite $B \subset A$ s.t. $\{E_\alpha\}_{\alpha \in B}$ is an open cover of $f(K)$.

First, I claim that $\{f^{-1}(E_\alpha)\}_{\alpha \in A}$ is an open cover of K.

Because each E_α is open in (Y, d_Y), the set $f^{-1}(E_\alpha)$ is open in (X, d_X). Take any $x \in K$. We have $f(x) \in f(K)$. Because $\{E_\alpha\}_{\alpha \in A}$ covers $f(K)$, there exists some $\hat{\alpha} \in A$ s.t. $f(x) \in E_{\hat{\alpha}}$. So $x \in f^{-1}(E_{\hat{\alpha}})$. Therefore, $\{f^{-1}(E_\alpha)\}_{\alpha \in A}$ is an open cover of K.

Because K is compact, there exists some finite set $B \subset A$ s.t. $\{f^{-1}(E_\alpha)\}_{\alpha \in B}$ covers K. Now, it is sufficient to show that $\{E_\alpha\}_{\alpha \in B}$ is an open cover of $f(K)$.

Take any $y \in f(K)$. There exists $\hat{x} \in K$ s.t. $f(\hat{x}) = y$. Because $\{f^{-1}(E_\alpha)\}_{\alpha \in B}$ covers K, there exists $\hat{\alpha} \in B$ s.t. $\hat{x} \in f^{-1}(E_{\hat{\alpha}})$. Therefore, $y = f(\hat{x}) \in E_{\hat{\alpha}}$. Therefore, $\{E_\alpha\}_{\alpha \in B}$ is an open cover of $f(K)$. $\square$

Although taking the image of a continuous function preserves compactness, the inverse image of a compact set may not be compact. For example, consider the function $f : \mathbb{R}_{++} \to \mathbb{R}$ defined as $g(x) := 1/x$. The inverse image of the compact set $[0, 1]$ under f is $[1, +\infty)$, which is not compact.

The next result states that a compact set in $(\mathbb{R}, d_2)$ has a maximum and a minimum.

Claim 5.13. *Let K be a compact set in $(\mathbb{R}, d_2)$. Then there exists $x^* \in K$ s.t. $x^* \geq x$ for any $x \in K$, and there exists $x_* \in K$ s.t. $x_* \leq x$ for any $x \in K$.*

Proof. Because K is compact in $(\mathbb{R}, d_2)$, we know that K is bounded, i.e. there exists some $B_r(x) \supset K$. Therefore, $x + r$ is an upper bound of K. By l.u.b. property of $\mathbb{R}$, there exists the least upper bound $\sup K$.

Now I claim that $\sup K \in K$.

Suppose $\sup K \notin K$. Then $\sup K$ is a limit point of K, because for any $\varepsilon > 0$, there exists $x \in K$ s.t. $x > \sup K - \varepsilon$. Because K is closed, we know that $K' \subset K$, and thus $\sup K \in K$. This contradicts the hypothesis we started with.

Let $x^* := \sup K$, and we have $x^* \in K$ and $x^* \geq x$ for any $x \in K$.

Symmetrically, let $x_* := \inf K$, and we can show that $x_* \in K$ and $x_* \leq x$ for any $x \in K$. □

Combining the two results above, we have Weierstrass theorem stated below.

Theorem 5.14 (Weierstrass)**.** *Let (X, d_X) be a metric space, and function $f : X \to \mathbb{R}$ is continuous. Let S be a compact set in (X, d_X). There exists $x^* \in S$ s.t. $f(x^*) \geq f(x)$ for any $x \in S$, and there exists $x_* \in S$ s.t. $f(x_*) \leq f(x)$ for any $x \in S$.*

Again, when we say $f : X \to \mathbb{R}$ is continuous, we use the Euclidean metric d_2 in the codomain $\mathbb{R}$ by default.

Proof. By Theorem 5.12, we know that $f(S)$ is compact in $(\mathbb{R}, d_2)$. Therefore, there exists $y^* \in f(S)$ s.t. $y^* \geq f(x)$ for any $x \in S$. By definition of the image $f(S)$, there exists $x^* \in S$ s.t. $f(x^*) = y^*$, and therefore $f(x^*) \geq f(x)$ for any $x \in S$.

Symmetrically, we can find x_*. □

In economics, it is standard to assume that every entity in the economy is maximizing some objective function. Weierstrass theorem implies that each entity's maximization problem must have a solution, if the entity's objective function is continuous and the set of alternatives available to the entity is compact.

6 Appendix

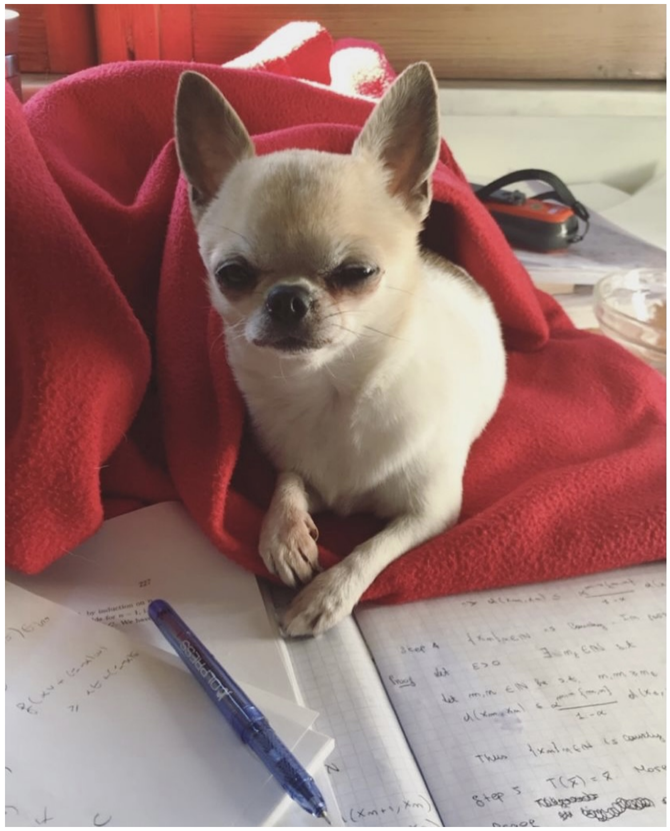

Figure 1: A picture of my dog, Leila, studying the proof of Banach Fixed Point theorem