# Chapter 3

# Analysis in Metric Space

Metric spaces are sets (spaces) with a notion of distance between points in the space that satisfies certain axioms. From these axioms we can deduce many properties relating to convergence, continuity, boundedness, and other concepts needed for the study of dynamics. Metric space theory provides both an elegant and powerful framework for analyzing the kinds of problems we wish to consider, and a great sandpit for playing with analytical ideas: A careful read of this chapter should strengthen your ability to read and write proofs.

The chapter supposes that you have at least some exposure to introductory real analysis or advanced calculus. A review of this material is given in appendix A. On the other hand, if you are already familiar with the fundamentals of metric spaces, then the best approach is to skim through this chapter quickly and return as necessary.

## 3.1 A First Look at Metric Space

Consider the set $\mathbb{R}^k$, a typical element of which is a *vector* $x = (x_1, \ldots, x_k)$, where $x_i \in \mathbb{R}$. There are several important topological notions we need to introduce for $\mathbb{R}^k$. These notions concern sets and functions on or into such space. In order to introduce them, it is convenient to begin with the concept of *Euclidean distance* between vectors, defined by

$$d_2(x, y) =: \|x - y\|_2 := \left[ \sum_{i=1}^{k} (x_i - y_i)^2 \right]^{1/2} \tag{3.1}$$

You have surely met this notion of distance before and you might know that it satisfies the following three conditions:

1. $d_2(x, y) = 0$ if and only if $x = y$,
2. $d_2(x, y) = d_2(y, x)$, and
3. $d_2(x, y) \leq d_2(x, v) + d_2(v, y)$.

for any $x, y, v \in \mathbb{R}^k$. The first property says that a point is at zero distance from itself, and also that distinct points always have positive distance. The second property is symmetry, and the third—the only one that is not immediately apparent—is the triangle inequality.

These three properties are fundamental to our understanding of distance. In fact if you look at the proofs of many important results—for example, the proof that every continuous function $f$ from a closed bounded subset of $\mathbb{R}^k$ to $\mathbb{R}$ has a maximizer and a minimizer—you will notice that no other properties of $d_2$ are actually used.

Now it turns out that there are many other "distance" functions we can impose on $\mathbb{R}^k$ that also satisfy properties 1–3. Any proof for the Euclidean (i.e., $d_2$) case that only uses properties 1–3 continues to hold for other distances, and in certain problems alternative notions of distance are easier to work with. This motivates us to generalize the concept of distance in $\mathbb{R}^k$.

While we are generalizing the notion of distance between vectors in $\mathbb{R}^k$, it is worth thinking about distance between other kinds of objects. If we could define the distance between two (infinite) sequences, or between a pair of functions, or two probability distributions, we could then give a definition for things like the "convergence" of distributions discussed informally in chapter 1.

### 3.1.1 Distances and Norms

Here is the key definition:

**Definition 3.1.1** A *metric space* is a nonempty set $S$ and a *metric* or *distance* $\rho \colon S \times S \to \mathbb{R}$ such that, for any $x, y, v \in S$,

1. $\rho(x, y) = 0$ if and only if $x = y$,
2. $\rho(x, y) = \rho(y, x)$, and
3. $\rho(x, y) \leq \rho(x, v) + \rho(v, y)$.

Apart from being nonempty, the set $S$ is completely arbitrary. In the context of a metric space the elements of the set are usually called points. As in the case of Euclidean distance, the third axiom is called the triangle inequality.

An immediate consequence of the axioms in definition 3.1.1 (which are sometimes referred to as the Hausdorff postulates) is that $\rho(x, y) \geq 0$ for any $x, y \in S$. To see this, note that if $x$ and $y$ are any two points in $S$, then $0 = \rho(x,x) \leq \rho(x,y) + \rho(y,x) = \rho(x,y) + \rho(x,y) = 2\rho(x,y)$. Hence $\rho(x,y) \geq 0$ as claimed.

The space $(\mathbb{R}^k, d_2)$ is a metric space, as discussed above. The most important case is $k = 1$, when $d_2(x,y)$ reduces to $|x - y|$ for $x, y \in \mathbb{R}$. *The notation* $(\mathbb{R}, |\cdot|)$ *will be used to denote this one-dimensional space.*

Many additional metric spaces on $\mathbb{R}^k$ are generated by what is known as a norm:

**Definition 3.1.2** A *norm* on $\mathbb{R}^k$ is a mapping $\mathbb{R}^k \ni x \mapsto \|x\| \in \mathbb{R}$ such that, for any $x, y \in \mathbb{R}^k$ and any $\gamma \in \mathbb{R}$,

1. $\|x\| = 0$ if and only if $x = 0$,
2. $\|\gamma x\| = |\gamma| \|x\|$, and
3. $\|x + y\| \leq \|x\| + \|y\|$.

Each norm $\|\cdot\|$ on $\mathbb{R}^k$ generates a metric $\rho$ on $\mathbb{R}^k$ via $\rho(x,y) := \|x - y\|$.

**Exercise 3.1** Verify the last claim by checking the axioms in definition 3.1.1.

**Exercise 3.2** Prove: $|\|x\| - \|y\|| \leq \|x - y\|$ for any norm $\|\cdot\|$ on $\mathbb{R}^k$ and $x, y \in \mathbb{R}^k$.

The most familiar norm on $\mathbb{R}^k$ is $\|x\|_2 := (\sum_{i=1}^k x_i^2)^{1/2}$, which generates the Euclidean distance $d_2$. A class of norms that includes $\|\cdot\|_2$ as a special case is the family $\|\cdot\|_p$ defined by

$$\|x\|_p := \left( \sum_{i=1}^k |x_i|^p \right)^{1/p} \qquad (x \in \mathbb{R}^k) \tag{3.2}$$

where $p \geq 1$. It is standard to admit $p = \infty$ in this family, with $\|x\|_\infty := \max_{1 \leq i \leq k} |x_i|$.

Proving that $\|\cdot\|_p$ is indeed a norm on $\mathbb{R}^k$ for arbitrary $p \geq 1$ is not difficult, but neither is it entirely trivial. In particular, establishing the triangle inequality (property 3 of the norm) requires the services of Minkowski's inequality. The latter is found in any text covering norms and is omitted.

**Exercise 3.3** Confirm that $\|\cdot\|_p$ is a norm on $\mathbb{R}^k$ for the cases $p = 1$ and $p = \infty$.

The class of norms $\|\cdot\|_p$ gives rise to the class of metric spaces $(\mathbb{R}^k, d_p)$, where $d_p(x,y) := \|x - y\|_p$ for all $x, y \in \mathbb{R}^k$.

So far all our spaces have involved different metrics on finite-dimensional vector space. Next let's consider an example of a "function space." Let $U$ be any set, let $bU$ be the collection of all bounded functions $f \colon U \to \mathbb{R}$ (i.e., $\sup_{x \in U} |f(x)| < \infty$), and let

$$d_\infty(f,g) =: \|f - g\|_\infty := \sup_{x \in U} |f(x) - g(x)| \tag{3.3}$$

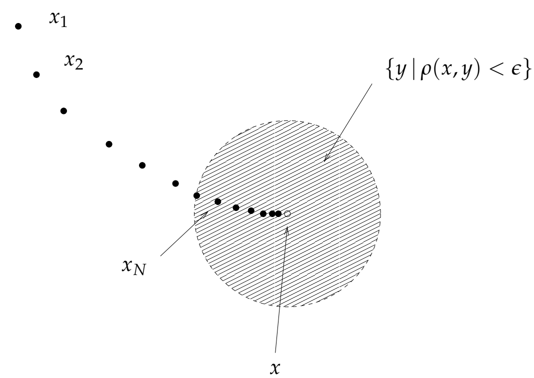


**Figure 3.1** Limit of a sequence

The space $(bU, d_\infty)$ is a metric space. Readers can check the first two properties of the definition of a metric space. The triangle inequality is verified as follows. Fix $f, g, h \in bU$ and $x \in U$. We have

$$|f(x) - g(x)| \leq |f(x) - h(x)| + |h(x) - g(x)| \leq d_\infty(f, h) + d_\infty(h, g)$$

Since $x$ is arbitrary, we obtain $d_\infty(f, g) \leq d_\infty(f, h) + d_\infty(h, g)$.[1]

### 3.1.2 Sequences

Let $S = (S, \rho)$ be a metric space. A sequence $(x_n) \subset S$ is said to *converge* to $x \in S$ if, for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that $n \geq N$ implies $\rho(x_n, x) < \epsilon$. In other words $(x_n)$ converges to $x$ if and only if the *real* sequence $\rho(x_n, x) \to 0$ in $\mathbb{R}$ as $n \to \infty$ (see §A.2 for more on real sequences). If this condition is satisfied, we write $\lim_{n\to\infty} x_n = x$, or $x_n \to x$. The point $x$ is referred to as the *limit* of the sequence. Figure 3.1 gives an illustration for the case of two-dimensional Euclidean space.

**Theorem 3.1.3** *A sequence in $(S, \rho)$ can have at most one limit.*

*Proof.* You might like to try a proof by contradiction as an exercise. Here is a direct proof. Let $(x_n)$ be an arbitrary sequence in $S$, and let $x$ and $x'$ be two limit points. We have

$$0 \leq \rho(x, x') \leq \rho(x, x_n) + \rho(x_n, x') \qquad \forall n \in \mathbb{N}$$

[1] As an aside, you may have noticed that the metric space $(bU, d_\infty)$ seems to be defined by a "norm" $\|f\|_\infty := \sup_{x\in U} |f(x)|$. This is not a norm in the sense of definition 3.1.2, as that definition requires that the underlying space is $\mathbb{R}^k$, rather than $bU$. However, more general norms can be defined for abstract "vector space," and $\|\cdot\|_\infty$ is a prime example. See, for example, Aliprantis and Burkinshaw (1998), Chapter 5.

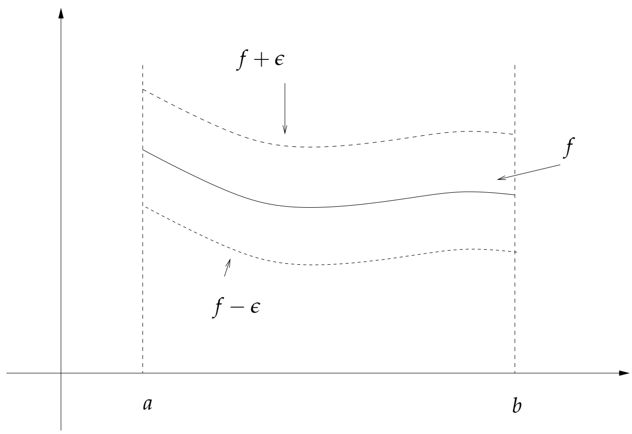


**Figure 3.2** An $\epsilon$-ball for $d_\infty$

From theorems A.2.8 and A.2.9 (page 332) we have $\rho(x, x') = 0$. Therefore $x = x'$. (Why?) □

**Exercise 3.4** Let $(x_n)$ and $(y_n)$ be sequences in $S$. Show that if $x_n \to x \in S$ and $\rho(x_n, y_n) \to 0$, then $y_n \to x$.

One of the most important creatures defined from the distance function is the humble open ball. The *open ball* or $\epsilon$-ball $B(\epsilon; x)$ centered on $x \in S$ with *radius* $\epsilon > 0$ is the set

$$B(\epsilon; x) := \{z \in S : \rho(z, x) < \epsilon\}$$

In the plane with $\rho = d_2$ the $\epsilon$-ball is a circle; in $\mathbb{R}^3$ it is a sphere. Figure 3.2 gives a visualization of the $\epsilon$-ball around $f \in (b[a, b], d_\infty)$.

**Exercise 3.5** Let $(x_n) \subset S$ and $x \in S$. Show that $x_n \to x$ if and only if for all $\epsilon > 0$, the ball $B(\epsilon; x)$ contains all but finitely many terms of $(x_n)$.

A subset $E$ of $S$ is called *bounded* if $E \subset B(n; x)$ for some $x \in S$ and some (suitably large) $n \in \mathbb{N}$. A sequence $(x_n)$ in $S$ is called bounded if its range $\{x_n : n \in \mathbb{N}\}$ is a bounded set.

**Exercise 3.6** Show that every convergent sequence in $S$ is also bounded.

Given sequence $(x_n) \subset S$, a subsequence is defined analogously to the case of real sequences: $(y_n)$ is called a *subsequence* of $(x_n)$ if there is a strictly increasing function

$f: \mathbb{N} \to \mathbb{N}$ such that $y_n = x_{f(n)}$ for all $n \in \mathbb{N}$. It is common to use the notation $(x_{n_k})$ to denote a subsequence of $(x_n)$.

**Exercise 3.7** Show that for $(x_n) \subset S$, $x_n \to x$ for some $x \in S$ if and only if every subsequence of $(x_n)$ converges to $x$.

For the Euclidean space $(\mathbb{R}^k, d_2)$ we have the following result:

**Lemma 3.1.4** *A sequence $(x_n) = (x_n^1, \ldots, x_n^k)$ in $(\mathbb{R}^k, d_2)$ converges to $x = (x^1, \ldots, x^k) \in \mathbb{R}^k$ if and only if $x_n^j \to x^j$ in $\mathbb{R} = (\mathbb{R}, |\cdot|)$ for all $j$ in $1, \ldots, k$.*

*Proof.* For $j$ in $1, \ldots, k$ we have $|x_n^j - x^j| \leq d_2(x_n, x)$. (Why?) Hence if $d_2(x_n, x) \to 0$, then $|x_n^j - x^j| \to 0$ for each $j$. For the converse, fix $\epsilon > 0$ and choose for each $j$ in $1, \ldots, k$ an $N^j \in \mathbb{N}$ such that $n \geq N^j$ implies $|x_n^j - x^j| < \epsilon/\sqrt{k}$. Now $n \geq \max_j N^j$ implies $d_2(x_n, x) \leq \epsilon$. (Why?) $\square$

Lemma 3.1.4 is important, and you should try sketching it for the case $k = 2$ to build intuition. We will see that in fact *the same result holds not just for $d_2$, but for the metric induced by any norm on $\mathbb{R}^k$.*

Let $S$ and $Y$ be two metric spaces. Parallel to §A.2.3, define $f: S \supset A \to Y$ to be *continuous* at $a \in A$ if for every sequence $(x_n)$ in $A$ converging to $a$ we have $f(x_n) \to f(a)$ in $Y$, and continuous on $A$ whenever it is continuous at every $a \in A$. For the same $f: A \to Y$ and for $a \in A$, we say that $y = \lim_{x \to a} f(x)$ if $f(x_n) \to y$ for every sequence $(x_n) \subset A$ with $x_n \to a$. Clearly, $f$ is continuous at $a$ if and only if $\lim_{x \to a} f(x) = f(a)$.

**Example 3.1.5** Let $S$ be a metric space, and let $\bar{x}$ be any given point in $S$. The map $S \ni x \mapsto \rho(x, \bar{x}) \in \mathbb{R}$ is continuous on all of $S$. To see this, pick any $x \in S$, and any $(x_n) \subset S$ with $x_n \to x$. Two applications of the triangle inequality yield

$$\rho(x, \bar{x}) - \rho(x_n, x) \leq \rho(x_n, \bar{x}) \leq \rho(x_n, x) + \rho(x, \bar{x}) \qquad \forall n \in \mathbb{N}$$

Now take the limit (i.e., apply theorem A.2.8 on page 332).

**Exercise 3.8** Let $f(x, y) = x^2 + y^2$. Show that $f$ is a continuous function from $(\mathbb{R}^2, d_2)$ into $(\mathbb{R}, |\cdot|)$.[2]

Throughout the text, if $S$ is some set, $f: S \to \mathbb{R}$, and $g: S \to \mathbb{R}$, then $f + g$ denotes the function $x \mapsto f(x) + g(x)$ on $S$, while $fg$ denotes the function $x \mapsto f(x)g(x)$ on $S$.

**Exercise 3.9** Let $f$ and $g$ be as above, and let $S$ be a metric space. Show that if $f$ and $g$ are continuous, then so are $f + g$ and $fg$.

---

[2]Hint: Use lemma 3.1.4.

**Exercise 3.10** A function $f\colon S \to \mathbb{R}$ is called upper-semicontinuous (usc) at $x \in S$ if, for every $x_n \to x$, we have $\limsup_n f(x_n) \leq f(x)$; and lower-semicontinuous (lsc) if, for every $x_n \to x$, we have $\liminf_n f(x_n) \geq f(x)$. Show that $f$ is usc at $x$ if and only if $-f$ is lsc at $x$. Show that $f$ is continuous at $x$ if and only if it is both usc and lsc at $x$.

### 3.1.3 Open Sets, Closed Sets

Arbitrary subsets of arbitrary spaces can be quite unruly. It is useful to identify classes of sets that are well-behaved, interacting nicely with common functions, and co-operating with attempts to measure them, or to represent them in terms of simpler elements. In this section we investigate a class of sets called the open sets, as well as their complements the closed sets.

Let's say that $x \in S$ *adheres* to $E \subset S$ if, for each $\epsilon > 0$, the ball $B(\epsilon; x)$ contains at least one point of $E$;[3] and that $x$ is *interior* to $E$ if $B(\epsilon; x) \subset E$ for some $\epsilon > 0$.[4] A set $E \subset S$ is called *open* if all points in $E$ are interior to $E$, and *closed* if $E$ contains all points that adhere to $E$. In the familiar metric space $(\mathbb{R}, |\cdot|)$, canonical examples are the intervals $(a, b)$ and $[a, b]$, which are open and closed respectively.[5] The concepts of open and closed sets turn out to be some of the most fruitful ideas in all of mathematics.

**Exercise 3.11** Show that a point in $S$ adheres to $E \subset S$ if and only if it is the limit of a sequence contained in $E$.

**Theorem 3.1.6** *A set $F \subset S$ is closed if and only if for every convergent sequence entirely contained in $F$, the limit of the sequence is also in $F$.*

*Proof.* Do the proof as an exercise if you can. If not, here goes. Suppose that $F$ is closed, and take a sequence in $F$ converging to some point $x \in S$. Then $x$ adheres to $F$ by exercise 3.11, and is therefore in $F$ by definition. Suppose, on the other hand, that the limit of every convergent sequence in $F$ belongs to $F$. Take any $x \in S$ that adheres to $F$. By exercise 3.11, there is a sequence in $F$ converging to it. Therefore $x \in F$, and $F$ is closed. $\square$

Open sets and closed sets are closely related. In fact we have the following fundamental theorem:

**Theorem 3.1.7** *A subset of an arbitrary metric space $S$ is open if and only if its complement is closed, and closed if and only if its complement is open.*

*Proof.* The proof is a good exercise. If you need a start, here is a proof that $G$ open implies $F := G^c$ closed. Take $(x_n) \subset F$ with $x_n \to x \in S$. We wish to show that $x \in F$.

---

[3] In some texts, $x$ is said to be a *contact point* of $E$.

[4] Try sketching some examples for the case of $(\mathbb{R}^2, d_2)$.

[5] If you find it hard to verify this now, you won't by the end of the chapter.

In fact this must be the case because, if $x \notin F$, then $x \in G$, in which case there is an $\epsilon > 0$ such that $B(\epsilon, x) \subset G$. (Why?) Such a situation is not possible when $(x_n) \subset F$ and $x_n \to x$. (Why?) □

We call $D(\epsilon; x) := \{z \in S : \rho(z, x) \leq \epsilon\}$ the *closed $\epsilon$-ball* centered on $x$. Every $D(\epsilon; x) \subset S$ is a closed set, as anticipated by the notation. To see this, take $(a_n) \subset D(\epsilon; x)$ converging to $a \in S$. We need to show that $a \in D(\epsilon; x)$ or, equivalently, that $\rho(a, x) \leq \epsilon$. Since $\rho(a_n, x) \leq \epsilon$ for all $n \in \mathbb{N}$, since limits preserve orders and since $y \mapsto \rho(y, x)$ is continuous, we have $\rho(a, x) = \lim \rho(a_n, x) \leq \epsilon$.

**Exercise 3.12** Likewise every open ball $B(\epsilon; x)$ in $S$ is an open set. Prove this directly, or repeat the steps of the previous example applied to $B(\epsilon; x)^c$.

You will not find it difficult to convince yourself that if $(S, \rho)$ is any metric space, then the whole set $S$ is itself both open and closed. (Just check the definitions carefully.) This can lead to some confusion. For example, suppose that we consider the metric space $(S, |\cdot|)$, where $S = (0, 1)$. Since $(0, 1)$ is the whole space, it is closed. At the same time, $(0, 1)$ is open as a subset of $(\mathbb{R}, |\cdot|)$. The properties of openness and closedness are relative rather than absolute.

**Exercise 3.13** Argue that for any metric space $(S, \rho)$, the empty set $\emptyset$ is both open and closed.

**Exercise 3.14** Show that if $(S, \rho)$ is an arbitrary metric space, and if $x \in S$, then the set $\{x\}$ is always closed.

**Theorem 3.1.8** *If $F$ is a closed, bounded subset of $(\mathbb{R}, |\cdot|)$, then $\sup F \in F$.*

*Proof.* Let $s := \sup F$. Since $F$ is closed it is sufficient to show there exists a sequence $(x_n) \subset F$ with $x_n \to s$. (Why?) By lemma A.2.13 (page 334) such a sequence exists. □

**Exercise 3.15** Prove that a sequence converges to a point $x$ if and only if the sequence is eventually in every open set containing $x$.

**Exercise 3.16** Prove: If $\{G_\alpha\}_{\alpha \in A}$ are all open, then so is $\cup_{\alpha \in A} G_\alpha$.

**Exercise 3.17** Show that if $A$ is finite and $\{G_\alpha\}_{\alpha \in A}$ is a collection of open sets, then $\cap_{\alpha \in A} G_\alpha$ is also open.

In other words, arbitrary unions and finite intersections of open sets are open. But be careful: An infinite intersection of open sets is not necessarily open. For example, consider the metric space $(\mathbb{R}, |\cdot|)$. If $G_n = (-1/n, 1/n)$, then $\cap_{n \in \mathbb{N}} G_n = \{0\}$ because

$$x \in \cap_n G_n \iff -\frac{1}{n} < x < \frac{1}{n} \quad \forall n \in \mathbb{N} \iff x = 0$$

**Exercise 3.18** Show that $\cap_{n \in \mathbb{N}}(a - 1/n, b + 1/n) = [a, b]$.

**Exercise 3.19** Prove that if $\{F_\alpha\}_{\alpha \in A}$ are all closed, then so is $\cap_{\alpha \in A} F_\alpha$.

**Exercise 3.20** Show that if $A$ is finite and $\{F_\alpha\}_{\alpha \in A}$ is a collection of closed sets, then $\cup_{\alpha \in A} F_\alpha$ is closed. On the other hand, show that the union $\cup_{n \in \mathbb{N}}[a + 1/n, b - 1/n] = (a, b)$. (Why is this not a contradiction?)

**Exercise 3.21** Show that $G \subset S$ is open if and only if it can be formed as the union of an arbitrary number of open balls.

**Remark 3.1.9** Later, when we try to make precise statements about dynamic systems evolving on some set $S \subset \mathbb{R}^n$, we will want $S$ to be a "nice" set, in some sense, to prevent the construction of strange and obscure counterexamples. We could assume that $S$ is open, since open sets are nice, but sometimes we want to work with closed sets. We could assume "either open or closed," but this also rules out some plausible scenarios (e.g., $S = [0, 1)$). Faced with this problem, we will typically assume that the set $S$ is a $G_\delta$ *set*, which means that $S$ can be expressed as a countable union of open sets. By constructions such as the one seen in exercise 3.18, we can represent every state space $S$ we care about in this text as a $G_\delta$ set. At the same time, elements of $G_\delta$ are regular enough to rule out most nasty counterexamples.

The *closure* of $E$ is the set of all points that adhere to $E$, and is written $\operatorname{cl} E$. In view of exercise 3.11, $x \in \operatorname{cl} E$ if and only if there exists a sequence $(x_n) \subset E$ with $x_n \to x$. The *interior* of $E$ is the set of its interior points, and is written $\operatorname{int} E$.

**Exercise 3.22** Show that $\operatorname{cl} E$ is always closed. Show in addition that for all closed sets $F$ such that $F \supset E$, $\operatorname{cl} E \subset F$. Using this result, show that $\operatorname{cl} E$ is equal to the intersection of all closed sets containing $E$.

The last exercise tells us that the closure of a set is the smallest closed set that contains that particular set. The next one shows us that the interior of a set is the largest open set contained in that set.

**Exercise 3.23** Show that $\operatorname{int} E$ is always open. Show also that for all open sets $G$ such that $G \subset E$, $\operatorname{int} E \supset G$. Using this result, show that $\operatorname{int} E$ is equal to the union of all open sets contained in $E$.

**Exercise 3.24** Show that $E = \operatorname{cl} E$ if and only if $E$ is closed. Show that $E = \operatorname{int} E$ if and only if $E$ is open.

Open sets and continuous functions interact very nicely. For example, we have the following fundamental theorem.

**Theorem 3.1.10** *A function $f\colon S \to Y$ is continuous if and only if the preimage $f^{-1}(G)$ of every open set $G \subset Y$ is open in $S$.*

*Proof.* Suppose that $f$ is continuous, and let $G$ be any open subset of $Y$. If $x \in f^{-1}(G)$, then $x$ must be interior, for if it is not, then there is a sequence $x_n \to x$ where $x_n \notin f^{-1}(G)$ for all $n$. But, by continuity, $f(x_n) \to f(x)$, implying that $f(x) \in G$ is not interior to $G$. (Why?) Contradiction.

Conversely, suppose that the preimage of every open set is open, and take any $\{x_n\}_{n\geq 1} \cup \{x\} \subset S$ with $x_n \to x$. Pick any $\epsilon$-ball $B$ around $f(x)$. The preimage $f^{-1}(B)$ is open, so for $N$ sufficiently large we have $x_n \in f^{-1}(B)$ for all $n \geq N$, in which case $f(x_n) \in B$ for all $n \geq N$. □

**Exercise 3.25** Let $S$, $Y$, and $Z$ be metric spaces, and let $f\colon S \to Y$ and $g\colon Y \to Z$. Show that if $f$ and $g$ are continuous, then so is $h := g \circ f$.

**Exercise 3.26** Let $S = \mathbb{R}^k$, and let $\rho^*(x,y) = 0$ if $x = y$ and 1 otherwise. Prove that $\rho^*$ is a metric on $\mathbb{R}^k$. Which subsets of this space are open? Which subsets are closed? What kind of functions $f\colon S \to \mathbb{R}$ are continuous? What kinds of sequences are convergent?

## 3.2 Further Properties

Having covered the fundamental ideas of convergence, continuity, open sets and closed sets, we now turn to two key concepts in metric space theory: completeness and compactness. After stating the definitions and covering basic properties, we will see how completeness and compactness relate to existence of optima and to the theory of fixed points.

### 3.2.1 Completeness

A sequence $(x_n)$ in metric space $(S,\rho)$ is said to be a *Cauchy* sequence if, for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that $\rho(x_j, x_k) < \epsilon$ whenever $j \geq N$ and $k \geq N$ . A subset $A$ of a metric space $S$ is called *complete* if every Cauchy sequence in $A$ converges to some point in $A$. Often the set $A$ of interest is the whole space $S$, in which case we say that $S$ is a complete metric space. As discussed in §A.2, the set of reals $(\mathbb{R}, |\cdot|)$ has this property. Many other metric spaces do not.

Notice that completeness is intrinsic to a given set $A$ and a metric $\rho$ on $A$. Either every Cauchy sequence in $(A,\rho)$ converges or there exists a Cauchy sequence that does not. On the other hand, openness and closedness are *relative* properties. The set $A := [0,1)$ is not open as a subset of $(\mathbb{R}, |\cdot|)$, but it is open as a subset of $(\mathbb{R}_+, |\cdot|)$.

The significance of completeness is that when searching for the solution to a problem, it sometimes happens that we are able to generate a Cauchy sequence whose limit would be a solution if it does in fact exist. In a complete space we can rest assured that our solution does exist as a well-defined element of the space.

**Exercise 3.27** Show that a sequence $(x_n)$ in metric space $(S, \rho)$ is Cauchy if and only if $\lim_{n\to\infty} \sup_{k\geq n} \rho(x_n, x_k) = 0$.

**Exercise 3.28** Show that if a sequence $(x_n)$ in metric space $(S, \rho)$ is convergent, then it is Cauchy. Show that if $(x_n)$ is Cauchy, then it is bounded.

Which metric spaces are complete? Observe that while $\mathbb{R} = (\mathbb{R}, |\cdot|)$ is complete, subsets of $\mathbb{R}$ may not be. For example, consider the metric space $(S, \rho) = ((0, \infty), |\cdot|)$. Some manipulation proves that while $(x_n) = (1/n)$ is Cauchy in $S$, it converges to no point in $S$. On the other hand, for $(S, \rho) = (\mathbb{R}_+, |\cdot|)$ the limit point of the sequence $(1/n)$ is in $S$. Indeed this space is complete, as is any closed subset of $\mathbb{R}$. More generally,

**Theorem 3.2.1** *Let $S$ be a complete metric space. Subset $A \subset S$ is complete if and only if it is closed as a subset of $S$.*

*Proof.* Let $A$ be complete. To see that $A$ is closed, let $(x_n) \subset A$ with $x_n \to x \in S$. Since $(x_n)$ is convergent it must be Cauchy (exercise 3.28). Because $A$ is complete we have $x \in A$. Thus $A$ contains its limit points, and is therefore closed. Conversely, suppose that $A$ is closed. Let $(x_n)$ be a Cauchy sequence in $A$. Since $S$ is complete, $(x_n)$ converges to some $x \in S$. As $A$ is closed, the limit point $x$ must be in $A$. Hence $A$ is complete. □

The Euclidean space $(\mathbb{R}^k, d_2)$ is complete. To see this, observe first that

**Lemma 3.2.2** *A sequence $(x_n) = (x_n^1, \ldots, x_n^k)$ in $(\mathbb{R}^k, d_2)$ is Cauchy if and only if each component sequence $x_n^j$ is Cauchy in $\mathbb{R} = (\mathbb{R}, |\cdot|)$.*

The proof of lemma 3.2.2 is an exercise.[6] The lemma is important because it implies that $(\mathbb{R}^k, d_2)$ inherits the completeness of $\mathbb{R}$ (axiom A.2.3, page 330):

**Theorem 3.2.3** *The Euclidean space $(\mathbb{R}^k, d_2)$ is complete.*

*Proof.* If $(x_n)$ is Cauchy in $(\mathbb{R}^k, d_2)$, then each component is Cauchy in $\mathbb{R} = (\mathbb{R}, |\cdot|)$, and, by completeness of $\mathbb{R}$, converges to some limit in $\mathbb{R}$. It follows from lemma 3.1.4 that $(x_n)$ is convergent in $(\mathbb{R}^k, d_2)$. □

[6] Hint: You might like to begin by rereading the proof of lemma 3.1.4.

Recall that $(bU, d_\infty)$ is the bounded real-valued functions $f\colon U \to \mathbb{R}$, endowed with the distance $d_\infty$ defined on page 41. This space also inherits completeness from $\mathbb{R}$:

**Theorem 3.2.4** *Let $U$ be any set. The metric space $(bU, d_\infty)$ is complete.*

*Proof.* Let $(f_n) \subset bU$ be Cauchy. We claim the existence of a $f \in bU$ such that $d_\infty(f_n, f) \to 0$. To see this, observe that for each $x \in U$ we have $\sup_{k \geq n} |f_n(x) - f_k(x)| \leq \sup_{k \geq n} d_\infty(f_n, f_k) \to 0$, and hence $(f_n(x))$ is Cauchy (see exercise 3.27). By the completeness of $\mathbb{R}$, $(f_n(x))$ is convergent, and we define a new function $f \in bU$ by $f(x) = \lim_{n\to\infty} f_n(x)$.[7]

To show that $d_\infty(f_n, f) \to 0$, fix $\epsilon > 0$, and choose $N \in \mathbb{N}$ such that $d_\infty(f_n, f_m) < \epsilon/2$ whenever $n, m \geq N$. Pick any $n \geq N$. For arbitrary $x \in U$ we have $|f_n(x) - f_m(x)| < \epsilon/2$ for all $m \geq n$, and hence, taking limits with respect to $m$, we have $|f_n(x) - f(x)| \leq \epsilon/2$. Since $x$ was arbitrary, $d_\infty(f_n, f) \leq \epsilon/2 < \epsilon$. □

This is a good opportunity to briefly discuss convergence of functions. A sequence of functions $(f_n)$ from arbitrary set $U$ into $\mathbb{R}$ converges *pointwise* to $f\colon U \to \mathbb{R}$ if $|f_n(x) - f(x)| \to 0$ as $n \to \infty$ for every $x \in U$; and *uniformly* if $d_\infty(f_n, f) \to 0$. Pointwise convergence is certainly important, but it is also significantly weaker than convergence in $d_\infty$. For example, suppose that $U$ is a metric space, that $f_n \to f$, and that all $f_n$ are continuous. It might then be hoped that the limit $f$ inherits continuity from the approximating sequence. For pointwise convergence this is not generally true,[8] while for uniform convergence it is:

**Theorem 3.2.5** *Let $(f_n)$ and $f$ be real-valued functions on metric space $U$. If $f_n$ is continuous on $U$ for all $n$ and $d_\infty(f_n, f) \to 0$, then $f$ is also continuous on $U$.*

*Proof.* Take $(x_k) \subset U$ with $x_k \to \bar{x} \in U$. Fix $\epsilon > 0$. Choose $n \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \epsilon/2$ for all $x \in U$. For any given $k \in \mathbb{N}$ the triangle inequality gives

$$|f(x_k) - f(\bar{x})| \leq |f(x_k) - f_n(x_k)| + |f_n(x_k) - f_n(\bar{x})| + |f_n(\bar{x}) - f(\bar{x})|$$

$$\therefore \quad |f(x_k) - f(\bar{x})| \leq |f_n(x_k) - f_n(\bar{x})| + \epsilon \qquad (k \in \mathbb{N})$$

From exercise A.20 (page 335) we have $0 \leq \limsup_k |f(x_k) - f(\bar{x})| \leq \epsilon$. Since $\epsilon$ is arbitrary, $\limsup_k |f(x_k) - f(\bar{x})| = \lim_k |f(x_k) - f(\bar{x})| = 0$. □

Now let's introduce another important metric space.

**Definition 3.2.6** Given any metric space $U$, let $(bcU, d_\infty)$ be the continuous functions in $bU$ endowed with the same metric $d_\infty$.

[7] Why is $f \in bU$ (i.e., why is $f$ bounded on $U$)? Consult exercise 3.28.

[8] A counterexample is $U = [0,1]$, $f_n(x) = x^n$, $f(x) = 0$ on $[0,1)$ and $f(1) = 1$.

**Theorem 3.2.7** *The space $(bcU, d_\infty)$ is always complete.*

*Proof.* This follows from theorem 3.2.1 (closed subsets of complete spaces are complete), theorem 3.2.4 (the space $(bU, d_\infty)$ is complete) and theorem 3.2.5 (which implies that the space $bcU$ is closed as a subset of $bU$). □

### 3.2.2 Compactness

Now we turn to the notion of compactness. A subset $K$ of $S = (S, \rho)$ is called *precompact* in $S$ if every sequence contained in $K$ has a subsequence that converges to a point of $S$. The set $K$ is called *compact* if every sequence contained in $K$ has a subsequence that converges to a point of $K$. (Thus every compact subset of $S$ is precompact in $S$, and every closed precompact set is compact.) Compactness will play a major role in our analysis. As we will see, the existence of a converging subsequence often allows us to track down the solution to a difficult problem.

As a first step, note that there is another important characterization of compactness, which at first sight bears little resemblance to the sequential definition above. To state the theorem, recall that for a set $K$ in $S$, an *open cover* is a collection $\{G_\alpha\}$ of open subsets of $S$ such that $K \subset \cup_\alpha G_\alpha$. The cover is called finite if it consists of only finitely many sets.

**Theorem 3.2.8** *A subset $K$ of an arbitrary metric space $S$ is compact if and only if every open cover of $K$ can be reduced to a finite cover.*

In other words, a set $K$ is compact if and only if, given any open cover, we can discard all but a finite number of elements and still cover $K$. The proof of theorem 3.2.8 can be found in any text on real analysis.

**Exercise 3.29** Exhibit an open cover of $\mathbb{R}^k$ that cannot be reduced to a finite subcover. Construct a sequence in $\mathbb{R}^k$ with no convergent subsequence.

**Exercise 3.30** Use theorem 3.2.8 to prove that every compact subset of a metric space $S$ is bounded (i.e., can be contained in an open ball $B(n; x)$ for some $x \in S$ and some (suitably large) $n \in \mathbb{N}$).

**Exercise 3.31** Prove that every compact subset of a metric space is closed.

On the other hand, closed and bounded subsets of metric spaces are not always compact.

**Exercise 3.32** Let $(S, \rho) = ((0, \infty), |\cdot|)$, and let $K = (0, 1]$. Show that although $K$ is a closed, bounded subset of $S$, it is not precompact in $S$.

**Exercise 3.33** Show that every subset of a compact set is precompact, and every closed subset of a compact set is compact.

**Exercise 3.34** Show that in any metric space the intersection of an arbitrary number of compact sets and the union of a finite number of compact sets are again compact.

**Exercise 3.35** For a more advanced exercise, you might like to try to show that the closure of a precompact set is compact. It follows that every precompact set is bounded. (Why?)

When it comes to precompactness and compactness, the space $(\mathbb{R}^k, d_2)$ is rather special. For example, the Bolzano–Weierstrass theorem states that

**Theorem 3.2.9** *Every bounded sequence in Euclidean space $(\mathbb{R}^k, d_2)$ has at least one convergent subsequence.*

*Proof.* Let's check the case $k = 2$. Let $(x_n) = (x_n^1, x_n^2) \subset (\mathbb{R}^2, d_2)$ be bounded. Since $(x_n^1)$ is itself bounded in $\mathbb{R}$ (why?), we can find a sequence $n_1, n_2, \ldots =: (n_j)$ such that $(x_{n_j}^1)$ converges in $\mathbb{R}$ (theorem A.2.6, page 332). Now consider $(x_{n_j}^2)$. This sequence is also bounded, and must itself have a convergent subsequence, so if we discard more terms from $n_1, n_2, \ldots =: (n_j)$ we can obtain a subsubsequence $(n_i) \subset (n_j)$ such that $(x_{n_i}^2)$ converges. Since $(n_i) \subset (n_j)$, the sequence $(x_{n_i}^1)$ also converges. It follows from lemma 3.1.4 (page 44) that $(x_{n_i})$ converges in $(\mathbb{R}^k, d_2)$. $\square$

The next result (called the Heine–Borel theorem) follows directly.

**Theorem 3.2.10** *A subset of $(\mathbb{R}^k, d_2)$ is precompact in $(\mathbb{R}^k, d_2)$ if and only if it is bounded, and compact if and only if it is closed and bounded.*

As we have seen, some properties of $(\mathbb{R}^k, d_2)$ carry over to arbitrary metric spaces, while others do not. For example, we saw that in an arbitrary metric space, closed and bounded sets are not necessarily compact. (This has important implications for the theory of Markov chains developed below.) However, we will see in §3.2.3 that any metric $d$ on $\mathbb{R}^k$ induced by a norm (see definition 3.1.2 on page 41) is "equivalent" to $d_2$, and that, as a result, subsets of $(\mathbb{R}^k, d)$ are compact if and only if they are closed and bounded.

### 3.2.3 Optimization, Equivalence

Optimization is important not only to economics, but also to statistics, numerical computation, engineering, and many other fields of science. In economics, rationality is the benchmark assumption for agent behavior, and is usually imposed by requiring agents solve optimization problems. In statistics, optimization is used for maximum likelihood and other related procedures, which search for the "best" estimator in some class. For numerical methods and approximation theory, one often seeks a simple representation $f_n$ of a given function $f$ that is the "closest" to $f$ in some suitable metric sense.

In any given optimization problem the first issue we must confront is whether or not optima exist. For example, a demand function is usually defined as the solution to a consumer optimization problem. It would be awkward then if no solution to the problem exists. The same can be said for supply functions, or for policy functions, which return the optimal action of a "controller" faced with a given state of the world. Discussions of existence typically begin with the following theorem:

**Theorem 3.2.11** *Let $f: S \to Y$, where $S$ and $Y$ are metric spaces and $f$ is continuous. If $K \subset S$ is compact, then so is $f(K)$, the image of $K$ under $f$.*

*Proof.* Take an open cover of $f(K)$. The preimage of this cover under $f$ is an open cover of $K$ (recall theorem 3.1.10 on page 48). Since $K$ is compact we can reduce this to a finite cover (theorem 3.2.8). The image of this finite cover under $f$ contains $f(K)$, and hence $f(K)$ is compact. □

**Exercise 3.36** Give another proof of theorem 3.2.11 using the sequential definitions of compactness and continuity.

The following theorem is one of the most fundamental results in optimization theory. It says that in the case of continuous functions on compact domains, optima always exist.

**Theorem 3.2.12** (Weierstrass) *Let $f: K \to \mathbb{R}$, where $K$ is a subset of arbitrary metric space $(S, \rho)$. If $f$ is continuous and $K$ is compact, then $f$ attains its supremum and infimum on $K$.*

In other words, $\alpha := \sup f(K)$ exists, and, moreover, there is an $x \in K$ such that $f(x) = \alpha$. A corresponding result holds for the infimum.

*Proof.* Regarding suprema, the result follows directly from theorem 3.2.11 combined with theorem 3.1.8 (page 46). By these theorems you should be able to show that $\alpha := \sup f(K)$ exists, and, moreover, that $\alpha \in f(K)$. By the definition of $f(K)$, there is an $x \in K$ such that $f(x) = \alpha$. This proves the assertion regarding suprema. The proof of the assertion regarding infima is similar. □

In general, for $f: S \to \mathbb{R}$, a value $y \in \mathbb{R}$ is called the *maximum* of $f$ on $A \subset S$ if $f(x) \leq y$ for all $x \in A$ and $f(\bar{x}) = y$ for some $\bar{x} \in A$. At most one maximum exists. The *maximizers* of $f$ on $A$ are the points

$$\operatorname*{argmax}_{x \in A} f(x) := \{x \in A : f(x) = y\} = \{x \in A : f(z) \leq f(x) \text{ for all } z \in A\}$$

Minima and minimizers are defined in a similar way.

With this notation, we can restate theorem 3.2.12 as follows: If $K$ is compact and $f\colon K \to \mathbb{R}$ is continuous, then $K$ contains at least one maximizer and one minimizer of $f$ on $K$. (Convince yourself that this is so.)

**Exercise 3.37** Let $f\colon K \to \mathbb{R}$, where $K$ is compact and $f$ is continuous. Show that if $f$ is strictly positive on $K$, then $\inf f(K)$ is strictly positive.

As an application of theorem 3.2.12, let's show that all norms on $\mathbb{R}^k$ induce essentially the same metric space. We begin with a definition: Let $S$ be a nonempty set, and let $\rho$ and $\rho'$ be two metrics on $S$. We say that $\rho$ and $\rho'$ are *equivalent* if there exist constants $K$ and $J$ such that

$$\rho(x,y) \leq K\rho'(x,y) \text{ and } \rho'(x,y) \leq J\rho(x,y) \qquad \text{for any } x,y \in S \tag{3.4}$$

The notion of equivalence is important because equivalent metrics share the same convergent sequences and Cauchy sequences, and the metric spaces $(S,\rho)$ and $(S,\rho')$ share the same open sets, closed sets, compact sets and bounded sets:

**Lemma 3.2.13** *Let $\rho$ and $\rho'$ be equivalent on $S$, and let $(x_n) \subset S$. The sequence $(x_n)$ $\rho$-converges to $x \in S$ if and only if it $\rho'$-converges to $x$.*[9]

*Proof.* If $\rho(x_n,x) \to 0$, then $\rho'(x_n,x) \leq J\rho(x_n,x) \to 0$, and so forth. $\square$

**Exercise 3.38** Let $\rho$ and $\rho'$ be equivalent on $S$, and let $(x_n) \subset S$. Show that $(x_n)$ is $\rho$-Cauchy if and only if it is $\rho'$-Cauchy.[10]

**Exercise 3.39** Let $\rho$ and $\rho'$ be equivalent on $S$, and let $A \subset S$. Show that $A$ is $\rho$-complete if and only if it is $\rho'$-complete.

**Exercise 3.40** Let $\rho$ and $\rho'$ be equivalent on $S$. Show that $(S,\rho)$ and $(S,\rho')$ share the same closed sets, open sets, bounded sets and compact sets.

**Exercise 3.41** Let $\rho$ and $\rho'$ be equivalent on $S$, and let $f\colon S \to \mathbb{R} = (\mathbb{R}, |\cdot|)$. Show that $f$ is $\rho$-continuous if and only if it is $\rho'$-continuous.

**Exercise 3.42** Let $S$ be any nonempty set, and let $\rho$, $\rho'$, and $\rho''$ be metrics on $S$. Show that equivalence is transitive, in the sense that if $\rho$ is equivalent to $\rho'$ and $\rho'$ is equivalent to $\rho''$, then $\rho$ is equivalent to $\rho''$.

**Theorem 3.2.14** *All metrics on $\mathbb{R}^k$ induced by a norm are equivalent.*

[9] Here $\rho$-convergence means convergence in $(S,\rho)$, etc., etc.

[10] Hint: Try a proof using exercise 3.27 (page 49).

*Proof.* The claim is that if $\|\cdot\|$ and $\|\cdot\|'$ are any two norms on $\mathbb{R}^k$ (see definition 3.1.2 on page 41), and $\rho$ and $\rho'$ are defined by $\rho(x,y) := \|x-y\|$ and $\rho'(x,y) := \|x-y\|'$, then $\rho$ and $\rho'$ are equivalent. In view of exercise 3.42, it is sufficient to show that any one of these metrics is equivalent to $d_1$. To check this, it is sufficient (why?) to show that if $\|\cdot\|$ is any norm on $\mathbb{R}^k$, then there exist constants $K$ and $J$ such that

$$\|x\| \leq K\|x\|_1 \text{ and } \|x\|_1 \leq J\|x\| \qquad \text{for any } x \in \mathbb{R}^k \tag{3.5}$$

To check the first inequality, let $e_j$ be the $j$-th basis vector in $\mathbb{R}^k$ (i.e., the $j$-th component of vector $e_j$ is 1 and all other components are zero). Let $K := \max_j \|e_j\|$. Then for any $x \in \mathbb{R}^k$ we have

$$\|x\| = \|x_1 e_1 + \cdots x_k e_k\| \leq \sum_{j=1}^{k} \|x_j e_j\| = \sum_{j=1}^{k} |x_j|\|e_j\| \leq K\|x\|_1$$

To check the second inequality in (3.5), observe that $x \mapsto \|x\|$ is continuous on $(\mathbb{R}^k, d_1)$ because if $x_n \to x$ in $d_1$, then

$$|\|x_n\| - \|x\|| \leq \|x_n - x\| \leq K\|x_n - x\|_1 \to 0 \quad (n \to \infty)$$

Now consider the set $E := \{x \in \mathbb{R}^k : \|x\|_1 = 1\}$. Some simple alterations to theorem 3.2.10 (page 52) and the results that lead to it show that, just as for the case of $(\mathbb{R}^k, d_2)$, closed and bounded subsets of $(\mathbb{R}^k, d_1)$ are compact.[11] Hence $E$ is $d_1$-compact. It now follows from theorem 3.2.12 that $x \mapsto \|x\|$ attains its minimum on $E$, in the sense that there is an $x^* \in E$ with $\|x^*\| \leq \|x\|$ for all $x \in E$. Clearly, $\|x^*\| \neq 0$. (Why?) Now observe that for any $x \in \mathbb{R}^k$ we have

$$\|x\| = \left\| \frac{x}{\|x\|_1} \right\| \|x\|_1 \geq \|x^*\| \|x\|_1$$

Setting $J := 1/\|x^*\|$ gives the desired inequality. □

### 3.2.4 Fixed Points

Next we turn to fixed points. Fixed point theory tells us how to find an $x$ that solves $Tx = x$ for some given $T \colon S \to S$.[12] Like optimization it has great practical importance. Very often the solutions of problems we study will turn out to be fixed points

[11] Alternatively, you can show directly that $(\mathbb{R}^k, d_2)$ and $(\mathbb{R}^k, d_1)$ are equivalent by establishing (3.5) for $\|\cdot\| = \|\cdot\|_2$. The first inequality is already done, and the second follows from the Cauchy–Schwartz inequality (look it up).

[12] It is common in fixed point theory to use upper case symbols like $T$ for the function, and no brackets around its argument. One reason is that $S$ is often a space of functions, and standard symbols like $f$ and $g$ are reserved for the elements of $S$.

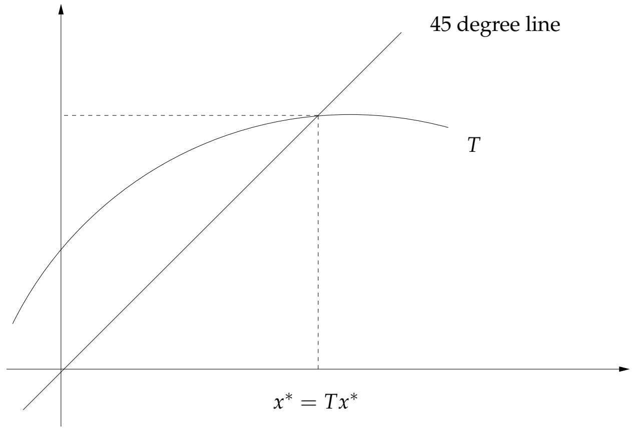


**Figure 3.3** Fixed points in one dimension

of some appropriately constructed function. Of the theorems we treat in this section, one uses convexity and is due to L. E. J. Brouwer while the other two are contraction mapping arguments: a famous one due to Stefan Banach and a variation of the latter.

Incidentally, fixed point and optimization problems are closely related. When we study dynamic programming, an optimization problem will be converted into a fixed point problem—in the process yielding an efficient means of computation. On the other hand, if $T\colon S \to S$ has a unique fixed point in metric space $(S, \rho)$, then finding that fixed point is equivalent to finding the minimizer of $g(x) := \rho(Tx, x)$.

So let $T\colon S \to S$, where $S$ is any set. An $x^* \in S$ is called a *fixed point* of $T$ on $S$ if $Tx^* = x^*$. If $S$ is a subset of $\mathbb{R}$, then fixed points of $T$ are those points in $S$ where $T$ meets the 45 degree line, as illustrated in figure 3.3.

**Exercise 3.43** Show that if $S = \mathbb{R}$ and $T\colon S \to S$ is decreasing ($x \leq y$ implies $Tx \geq Ty$), then $T$ has at most one fixed point.

A set $S \subset \mathbb{R}^k$ is called *convex* if for all $\lambda \in [0,1]$ and $a, a' \in S$ we have $\lambda a + (1-\lambda)a' \in S$. Here is Brouwer's fixed point theorem:

**Theorem 3.2.15** (Brouwer) *Consider the space $(\mathbb{R}^k, d)$, where $d$ is the metric induced by any norm.*[13] *Let $S \subset \mathbb{R}^k$, and let $T\colon S \to S$. If $T$ is continuous and $S$ is both compact and convex, then $T$ has at least one fixed point in $S$.*

The proof is rather long and we omit it. I recommend you sketch the case $S = [0,1]$ to gain some intuition.

[13] All such metrics are equivalent. See theorem 3.2.14.

Let $(S, \rho)$ be a metric space, and let $T \colon S \to S$. The map $T$ is called *nonexpansive* on $S$ if

$$\rho(Tx, Ty) \leq \rho(x, y) \qquad \forall\, x, y \in S \tag{3.6}$$

It is called *contracting* on $S$ if

$$\rho(Tx, Ty) < \rho(x, y) \qquad \forall\, x, y \in S \text{ with } x \neq y \tag{3.7}$$

It is called *uniformly contracting* on $S$ with modulus $\lambda$ if $0 \leq \lambda < 1$ and

$$\rho(Tx, Ty) \leq \lambda \rho(x, y) \qquad \forall\, x, y \in S \tag{3.8}$$

**Exercise 3.44** Show that if $T$ is nonexpansive on $S$ then it is also continuous on $S$ (with respect to the same metric $\rho$).

**Exercise 3.45** Show that if $T$ is a contraction on $S$, then $T$ has at most one fixed point in $S$.

For $n \in \mathbb{N}$ the notation $T^n$ refers to the $n$-th composition of $T$ with itself, so $T^n x$ means apply $T$ to $x$, apply $T$ to the result, and so on for $n$ times. By convention, $T^0$ is the identity map $x \mapsto x$.[14]

**Exercise 3.46** Let $T$ be uniformly contracting on $S$ with modulus $\lambda$, and let $x_0 \in S$. Define $x_n := T^n x_0$ for $n \in \mathbb{N}$. Use induction to show that $\rho(x_{n+1}, x_n) \leq \lambda^n \rho(x_1, x_0)$ for all $n \in \mathbb{N}$.

The next theorem is one of the cornerstones of functional analysis:

**Theorem 3.2.16** (Banach) *Let $T \colon S \to S$, where $(S, \rho)$ is a complete metric space. If $T$ is a uniform contraction on $S$ with modulus $\lambda$, then $T$ has a unique fixed point $x^* \in S$. Moreover for every $x \in S$ and $n \in \mathbb{N}$ we have $\rho(T^n x, x^*) \leq \lambda^n \rho(x, x^*)$, and hence $T^n x \to x^*$ as $n \to \infty$.*

*Proof.* Let $\lambda$ be as in (3.8). Let $x_n := T^n x_0$, where $x_0$ is some point in $S$. From exercise 3.46 we have $\rho(x_n, x_{n+1}) \leq \lambda^n \rho(x_0, x_1)$ for all $n \in \mathbb{N}$, suggesting that the sequence is $\rho$-Cauchy. In fact with a bit of extra work one can show that if $n, k \in \mathbb{N}$ and $n < k$, then $\rho(x_n, x_k) \leq \sum_{i=n}^{k-1} \lambda^i \rho(x_0, x_1)$.

$$\therefore \quad \rho(x_n, x_k) < \frac{\lambda^n}{1 - \lambda} \rho(x_0, x_1) \qquad (n, k \in \mathbb{N} \text{ with } n < k)$$

Since $(x_n)$ is $\rho$-Cauchy, this sequence has a limit $x^* \in S$. That is, $T^n x_0 \to x^* \in S$. Next we show that $x^*$ is a fixed point of $T$. Since $T$ is continuous, we have $T(T^n x_0) \to Tx^*$.

[14] In other words, $T^0 := \{x \mapsto x\}$ and $T^n := T \circ T^{n-1}$ for $n \in \mathbb{N}$.

But $T(T^n x_0) \to x^*$ clearly also holds. (Why?) Since sequences in a metric space have at most one limit, it must be that $Tx^* = x^*$.

Regarding uniqueness, let $x$ and $x'$ be fixed points of $T$ in $S$. Then

$$\rho(x, x') = \rho(Tx, Tx') \leq \lambda \rho(x, x')$$

$$\therefore \quad \rho(x, x') = 0, \text{ and hence } x = x'$$

The estimate $\rho(T^n x, x^*) \leq \lambda^n \rho(x, x^*)$ in the statement of the theorem is left as an exercise. □

If we take away uniformity and just have a contraction, then Banach's proof of stability does not work, and indeed a fixed point may fail to exist. Under the action of a *uniformly* contracting map $T$, the motion induced by iterating $T$ slows down at a geometric rate. The limit of this process is a fixed point. On the other hand, with a contraction we know only that the process slows down at each step, and this is not enough to guarantee convergence. Imagine a particle that travels at speed $1 + 1/t$ at time $t$. Its motion slows down at each step, but the particle's speed is bounded away from zero.

**Exercise 3.47** Let $S := \mathbb{R}_+$ with distance $|\cdot|$, and let $T: x \mapsto x + e^{-x}$. Show that $T$ is a contraction on $S$, and that $T$ has no fixed point in $S$.

However, if we add compactness of $S$ to the contractiveness of $T$ the problem is rectified. Now our particle cannot diverge, as that would violate the existence of a convergent subsequence.

**Theorem 3.2.17** *If $(S, \rho)$ is compact and $T: S \to S$ is contracting, then $T$ has a unique fixed point $x^* \in S$. Moreover $T^n x \to x^*$ for all $x \in S$.*

The proof is provided in the appendix to this chapter (p. 343).

## 3.3 Commentary

The French mathematician Maurice Fréchet (1878–1973) introduced the notion of metric space in his dissertation of 1906. The name "metric space" is due to Felix Hausdorff (1868–1942). Other important spaces related to metric spaces are topological spaces (a generalization of metric space) and normed linear spaces (metric spaces with additional algebraic structure). Good references on metric space theory—sorted from elementary to advanced—include Light (1990), Kolmogorov and Fomin (1970), Aliprantis and Burkinshaw (1998), and Aliprantis and Border (1999). For a treatment with economic applications, see Ok (2007).

This chapter's discussion of fixed points and optimization only touched the surface of these topics. For a nice treatment of optimization theory, see Sundaram (1996). Various extensions of Brouwer's fixed point theorem are available, including Kakutani's theorem (for correspondences, see McLennan and Tourky 2005 for an interesting proof) and Schauder's theorem (for infinite-dimensional spaces). Aliprantis and Border (1999) is a good place to learn more. See Aguiar and Amador (2019) for a creative use of contraction maps in the setting of sovereign debt models.