Terence Tao

Analysis II (Exercise)

Third Edition

Contents

CHAPTERI

Metric spaces

Definitions and examples

Exercise 1.1.1. *Prove Lemma 1.1.1.*

> **Lemma 1.1.1.** Let $(x_n)_{n=m}^{\infty}$ be a sequence of real numbers, and let x be another real number. Then $(x_n)_{n=m}^{\infty}$ converges to x if and only if $\lim_{n\to\infty} d(x_n, x) = 0$.

Proof. Since $(x_n)_{n=m}^{\infty}$ converges to x, by Theorem 6.1.19(d),

$$\lim_{n\to\infty} d(x_n, x) = \lim_{n\to\infty} |x_n - x| = \left| \lim_{n\to\infty} x_n - x \right| = |x - x| = 0.$$

Conversely, we have $\lim_{n\to\infty} x_n = x$. Thus $(x_n)_{n=m}^{\infty}$ converges to x if and only if $\lim_{n\to\infty} d(x_n, x) = 0$. $\qquad\square$

Exercise 1.1.2. *Show that the real line with the metric $d(x, y) := |x - y|$ is indeed a metric space. (Hint: you may wish to review your proof of Proposition 4.3.3.)*

Proof. By Proposition 4.3.3(e), (f) and (g), the metric $d(x, y) := |x - y|$ is

a metric space by Definition 1.1.2.   □

Exercise 1.1.3. *Let X be a set, and let $d : X \times X \to [0, \infty)$ be a function.*

(a) *Give an example of a pair (X, d) which obeys axioms (bcd) of Definition 1.1.2, but not (a). (Hint: modify the discrete metric.)*

(b) *Give an example of a pair (X, d) which obeys axioms (acd) of Definition 1.1.2, but not (b).*

(c) *Give an example of a pair (X, d) which obeys axioms (abd) of Definition 1.1.2, but not (c).*

(d) Give an example of a pair (X, d) which obeys axioms (abc) of Definition 1.1.2, but not (d). (Hint: try examples where X is a finite set.)

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

(a) Define the metric d by setting $d(x, y) := 1$ for all $x, y \in X$.

(b) Define the metric d by setting $d(x, y) := 0$ for all $x, y \in X$.

(c) Define the metric d by setting $d(x, y) := 0$ when $x = y$, and $d(x, y) = x - y$ when $x \neq y$.

(d) Let $X := \{x, y, z\}$. Define the metric d by setting $d(x, x) = d(y, y) = d(z, z) = 0$, $d(x, y) = d(y, x) = 1$, $d(x, z) = d(z, x) = 3$, and $d(y, z) = d(z, y) = 1$.   □

Exercise 1.1.4. *Show that the pair $(Y, d|_{Y \times Y})$ defined in Example 1.1.5 is indeed a metric space.*

Proof. Let (X, d) be any metric space, and let Y be a subset of X. Then we can restrict the metric function $d : X \times X \to [0, +\infty)$ to the subset $Y \times Y$ of $X \times X$ to create a restricted metric function $d|_{Y \times Y} : Y \times Y \to [0, +\infty)$ of Y; this is known as the metric on Y *induced* by the metric d on X.

For any $x, y, z \in Y \times Y$, we have $x, y, z \in X \times X$ for that $Y \times Y \subseteq X \times X$. Then by Definition 1.1.2, $(Y, d|_{Y \times Y})$ is a metric space.   □

Exercise 1.1.5. *Let $n \geq 1$, and let $a_1, a_2, \cdots, a_n$ and $b_1, b_2, \cdots, b_n$ be real numbers. Verify the identity*

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

and conclude the Cauchy-Schwarz inequality

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

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

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

Proof. We use induction on n. When $n = 1$, equation is hold for that

$$(a_1 b_1)^2 + \frac{1}{2}(a_1 b_1 - a_1 b_1)^2 = a_1^2 b_1^2.$$

Now we inductively suppose that

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

is hold. Since

$$\left(\sum_{i=1}^{n+1} a_i b_i \right)^2 = \left(\sum_{i=1}^{n} a_i b_i + a_{n+1} b_{n+1} \right)^2$$

$$= \left(\sum_{i=1}^{n} a_i b_i \right)^2 + 2 a_{n+1} b_{n+1} \left(\sum_{i=1}^{n} a_i b_i \right) + (a_{n+1} b_{n+1})^2,$$

and

$$\sum_{i=1}^{n+1} \sum_{j=1}^{n+1} (a_i b_j - a_j b_i)^2 = \sum_{i=1}^{n+1} \left(\sum_{j=1}^{n} (a_i b_j - a_j b_i)^2 + (a_i b_{n+1} - a_{n+1} b_i)^2 \right)$$

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

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

where
$$\sum_{i=1}^{n}(a_ib_{n+1} - a_{n+1}b_i)^2 = \sum_{i=1}^{n}\left(a_i^2 b_{n=1}^2 + a_{n+1}^2 b_i^2 - 2a_{n+1}b_{n+1}a_ib_i\right)$$

$$= b_{n+1}^2 \sum_{i=1}^{n} a_i^2 + a_{n+1}^2 \sum_{i=1}^{n} b_i^2 - 2a_{n+1}b_{n+1}\left(\sum_{i=1}^{n} a_ib_i\right).$$

Thus
$$\left(\sum_{i=1}^{n+1} a_ib_i\right)^2 + \frac{1}{2}\sum_{i=1}^{n+1}\sum_{j=1}^{n+1}(a_ib_j - a_jb_i)^2$$

$$= \left(\sum_{i=1}^{n} a_ib_i\right)^2 + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_ib_j - a_jb_i)^2$$

$$+ 2a_{n+1}b_{n+1}\left(\sum_{i=1}^{n} a_ib_i\right) + (a_{n+1}b_{n+1})^2 + \sum_{i=1}^{n}(a_ib_{n+1} - a_{n+1}b_i)^2$$

$$= \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{j=1}^{n} b_j^2\right) + 2a_{n+1}b_{n+1}\left(\sum_{i=1}^{n} a_{\sqrt{i}}b_i\right) + (a_{n+1}b_{n+1})^2$$

$$+ b_{n+1}^2 \sum_{i=1}^{n} a_i^2 + a_{n+1}^2 \sum_{i=1}^{n} b_i^2 - 2a_{n+1}b_{n+1}\left(\sum_{i=1}^{n} a_ib_i\right)$$

$$= \left(\sum_{i=1}^{n} a_i^2 + a_{n+1}^2\right)\left(\sum_{j=1}^{n} b_j^2 + b_{n+1}^2\right) = \left(\sum_{i=1}^{n+1} a_i^2\right)\left(\sum_{j=1}^{n+1} b_j^2\right).$$

This close the induction. Therefore, the identity is hold for all $n \geq 1$. In particular, since
$$\left(\sum_{i=1}^{n} a_ib_i\right)^2 \leq \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{j=1}^{n} b_j^2\right),$$

we have the Cauchy-Schwarz inequality
$$\left|\sum_{i=1}^{n} a_ib_i\right| \leq \left(\sum_{i=1}^{n} a_i^2\right)^{1/2}\left(\sum_{j=1}^{n} b_j^2\right)^{1/2}.$$

From the identity
$$\sum_{n=1}^{n}(a_i + b_i)^2 = \sum_{i=1}^{n} a_i^2 + \sum_{i=1}^{n} b_i^2 + 2\sum_{i=1}^{n} a_ib_i$$

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

$$= \left(\left(\sum_{i=1}^{n} a_i^2 \right)^{1/2} + \left(\sum_{i=1}^{n} b_i^2 \right)^{1/2} \right)^2.$$

Hence

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

The triangle inequality is hold. $\square$

Exercise 1.1.6. *Show that $(\mathbf{R}^n, d_{l^2})$ in Example 1.1.6 is indeed a metric space. (Hint: use Exercise 1.1.5.)*

Proof. Let $n \geq 1$ be a natural number, and let $\mathbf{R}^n$ be the space of n-tuples of real numbers:

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

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

$$d_{l^2}((x_1, \cdots, x_n), (y_1, \cdots, y_n)) := \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$$

$$= \left(\sum_{i=1}^{n} (x_i - y_i)^2 \right)^{1/2}.$$

Let $x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n), z = (z_1, \cdots, z_n)$

(a) For all $x \in \mathbf{R}^n$, if $x = x$, then $d_{l^2}(x, x) = 0$.

(b) For any distinct $x, y \in \mathbf{R}^n$, we have

$$d_{l^2}(x, y) = \left(\sum_{i=1}^{n} (x_i - y_i)^2 \right)^{1/2} \geq 0.$$

(c) For any $x, y \in \mathbf{R}^n$, we have

$$\left(\sum_{i=1}^{n} (x_i - y_i)^2 \right)^{1/2} = \left(\sum_{i=1}^{n} (y_i - x_i)^2 \right)^{1/2}$$

So that $d_{l^2}(x, y) = d_{l^2}(y, x)$.

(d) For any $x, y, z \in \mathbf{R}^n$, by triangle inequality, we have

$$\left(\sum_{i=1}^{n}(x_i - z_i)^2\right)^{1/2} = \left(\sum_{i=1}^{n}((x_i - y_i) + (y_i - z_i))^2\right)^{1/2}$$

$$\leq \left(\sum_{i=1}^{n}(x_i - y_i)^2\right)^{1/2} + \left(\sum_{i=1}^{n}(y_i - z_i)^2\right)^{1/2}.$$

So that $d_{l^2}(x, z) \leq d_{l^2}(x, y) + d_{l^2}(y, z)$.

Thus $(\mathbf{R}^n, d_{l^2})$ is a metric space.   □

Exercise 1.1.7. *Show that the pair $(\mathbf{R}^n, d_{l^1})$ in Example 1.1.7 is indeed a metric space.*

Proof. Let $n \geq 1$ be a natural number, and let $\mathbf{R}^n$ be the space of n-tuples of real numbers:

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

We define the *Taxi-cab metric* (or l^1 metric) $d_{l^1} : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ by

$$d_{l^1}((x_1, \cdots, x_n), (y_1, \cdots, y_n)) := |x_1 - y_1| + \cdots + |x_n - y_n|$$

$$= \sum_{i=1}^{n}|x_i - y_i|.$$

Let $x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n), z = (z_1, \cdots, z_n)$.

(a) For all $x \in \mathbf{R}^n$, if $x = x$, then $d_{l^1}(x, x) = 0$.

(b) For any distinct $x, y \in \mathbf{R}^n$, we have

$$d_{l^1}(x, y) = \sum_{i=1}^{n}|x_i - y_i| \geq 0.$$

(c) For any $x, y \in \mathbf{R}^n$, we have

$$\sum_{i=1}^{n}|x_i - y_i| = \sum_{i=1}^{n}|y_i - x_i|$$

So that $d_{l^1}(x, y) = d_{l^1}(y, x)$.

(d) For any $x, y, z \in \mathbf{R}^n$, by triangle inequality, we have

$$\sum_{i=1}^{n}|x_i - z_i| = \sum_{i=1}^{n}|x_i - y_i| + |y_i - z_i| \leq \sum_{i=1}^{n}|x_i - y_i| + \sum_{i=1}^{n}|y_i - z_i|.$$

So that $d_{l^1}(x, z) \le d_{l^1}(x, y) + d_{l^1}(y, z)$.

Thus $(\mathbf{R}^n, d_{l^1})$ is a metric space. $\qquad \square$

Exercise 1.1.8. *Prove the two inequalities in (1.1). (For the first inequality, square both sides. For the second inequality, use Exercise (1.1.5).)*

Proof. Let $x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n)$. Prove the inequalities

$$d_{l^2}(x, y) \le d_{l^1}(x, y) \le \sqrt{n} d_{l^2}(x, y).$$

For the first inequality, square both sides, then

$$\sum_{i=1}^n (x_i - y_i)^2 \le \left(\sum_{i=1}^n |x_i - y_i| \right)^2.$$

By the Cauchy-Schwarz inequality, the first inequality is hold. For the second inequality, we have

$$\sum_{i=1}^n |x_i - y_i| \le \left(\sum_{i=1}^n (x_i - y_i)^2 \right)^{1/2} \left(\sum_{i=1}^n 1^2 \right)^{1/2} = \sqrt{n} \left(\sum_{i=1}^n (x_i - y_i)^2 \right)^{1/2}.$$

Therefore $d_{l^2}(x, y) \le d_{l^1}(x, y) \le \sqrt{n} d_{l^2}(x, y)$ is hold for all x, y. $\qquad \square$

Exercise 1.1.9. *Show that the pair $(\mathbf{R}^n, d_{l^\infty})$ in Example 1.1.9 is indeed a metric space.*

Proof. Let $n \ge 1$ be a natural number, and let $\mathbf{R}^n$ be the space of n-tuples of real numbers:

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

We define the *sup norm metric* (or l^∞ metric) $d_{l^\infty} : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ by

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

Let $x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n), z = (z_1, \cdots, z_n)$

(a) For all $x \in \mathbf{R}^n$, if $x = x$, then $d_{l^\infty}(x, x) = \sup\{0 : 1 \le i \le n\} = 0$.

(b) For any distinct $x, y \in \mathbf{R}^n$, we have

$$d_{l^\infty}(x, y) = \sup\{|x_i - y_i| : 1 \le i \le n\} \ge |x_i - y_i| > 0.$$

(c) For any $x, y \in \mathbf{R}^n$, we have $|x_i - y_i| = |y_i - x_i|$, then $\sup\{|x_i - y_i| : 1 \le i \le n\} = \sup\{|y_i - x_i| : 1 \le i \le n\}$. So that $d_{l^\infty}(x, y) = d_{l^\infty}(y, x)$.

(d) For any $x, y, z \in \mathbf{R}^n$. Since $|x_i - z_i| \leq |x_i - y_i| + |y_i - z_i|$ for all $1 \leq i \leq n$, we have

$$|x_i - z_i| \leq d_{l^\infty}(x, z) \leq |x_i - y_i| + |y_i - z_i| \leq d_{l^\infty}(x, y) + d_{l^\infty}(y, z).$$

So that $d_{l^\infty}(x, z) \leq d_{l^\infty}(x, y) + d_{l^\infty}(y, z)$.

Thus $(\mathbf{R}^n, d_{l^1})$ is a metric space. □

Exercise 1.1.10. *Prove the two inequalities in (1.2).*

Proof. Let $x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_n)$. Prove the inequalities

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

Suppose that there is a $1 \leq m \leq n$ such that $|x_m - y_m| \geq |x_i - y_i|$ for all $1 \leq i \leq n$. Then we have

$$(\sup\{|x_i - y_i| : 1 \leq i \leq n\})^2 = |x_m - y_m|^2 = \frac{1}{n} \sum_{i=1}^{n} (x_m - y_m)^2.$$

Since

$$\frac{1}{n} \sum_{i=1}^{n} (x_i - y_i)^2 \leq \frac{1}{n} \sum_{i=1}^{n} (x_m - y_m)^2.$$

Then

$$\frac{1}{\sqrt{n}} \left(\sum_{i=1}^{n} (x_i - y_i)^2 \right)^{1/2} \leq \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_m - y_m)^2} = |x_m - y_m|.$$

This is the first inequality. For the second inequality, we have

$$|x_m - y_m| = \sqrt{(x_m - y_m)^2} \leq \left(\sum_{i=1}^{n} (x_i - y_i)^2 \right)^{1/2}$$

Thus $(1/\sqrt{n}) d_{l^2}(x, y) \leq d_{l^\infty}(x, y) \leq d_{l^2}(x, y)$ is hold for all x, y. □

Exercise 1.1.11. *Show that the discrete metric $(\mathbf{R}^n, d_{\mathrm{disc}})$ in Example 1.1.11 is indeed a metric space.*

Proof. Let X be an arbitrary set (finite or infinite), and define the *discrete metric* d_{disc} by setting $d_{\mathrm{disc}}(x, y) := 0$ when $x = y$, and $d_{\mathrm{disc}}(x, y) := 1$ when $x \neq y$.

(*a*) For all $x \in X$, if $x = x$, then $d_{\text{disc}}(x, x) = 0$.

(*b*) For any distinct $x, y \in X$, we have $d_{\text{disc}}(x, y) = 1 > 0$.

(*c*) For any $x, y \in X$, if $x = y$ then $d_{\text{disc}}(x, y) = d_{\text{disc}}(y, x) = 0$; while if $x \neq y$, then $d_{\text{disc}}(x, y) = d_{\text{disc}}(y, x) = 1$.

(*d*) For any $x, y, z \in X$.

If $x = z = y$ then $d_{\text{disc}}(x, z) = d_{\text{disc}}(x, y) + d_{\text{disc}}(y, z)$.

If $x = z \neq y$ then $d_{\text{disc}}(x, z) \leq d_{\text{disc}}(x, y) + d_{\text{disc}}(y, z)$.

If $x \neq z = y$, then $d_{\text{disc}}(x, z) = d_{\text{disc}}(x, y) + d_{\text{disc}}(y, z)$.

If $x = y \neq z$, then $d_{\text{disc}}(x, z) = d_{\text{disc}}(x, y) + d_{\text{disc}}(y, z)$.

If $x \neq z \neq y$ and $x \neq z$, then $d_{\text{disc}}(x, z) \leq d_{\text{disc}}(x, y) + d_{\text{disc}}(y, z)$.

Thus $d_{\text{disc}}(x, z) \leq d_{\text{disc}}(x, y) + d_{\text{disc}}(y, z)$.

Thus (X, d_{disc}) is a metric space. $\qquad\square$

Exercise 1.1.12. *Prove Proposition 1.1.18.*

> **Proposition 1.1.18** (Equivalence of l_1, l_2, l_∞). Let $\mathbf{R}^n$ be a Euclidean space, and let $(x^{(k)})_{k=m}^\infty$ be a sequence of points in $\mathbf{R}^n$. We write $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})$, i.e., for $j = 1, 2, \cdots, n$, $x_j^{(k)} \in \mathbf{R}$ is the jth coordinate of $x^{(k)} \in \mathbf{R}^n$. Let $x = (x_1, \cdots, x_n)$ be a point in $\mathbf{R}^n$. Then the following four statements are equivalent:
>
> (*a*) $(x^{(k)})_{k=m}^\infty$ converges to x with respect to the Euclidean metric d_{l_2}.
>
> (*b*) $(x^{(k)})_{k=m}^\infty$ converges to x with respect to the taxi-cab metric d_{l_1}.
>
> (*c*) $(x^{(k)})_{k=m}^\infty$ converges to x with respect to the sup norm metric d_{l_∞}.
>
> (*d*) For every $1 \leq j \leq n$, the sequence $(x^{(k)})_{k=m}^\infty$ converges to x_j. (Notice that this is a sequence of real numbers, not of points in $\mathbf{R}^n$.)

Proof. (a) $\Rightarrow$ (b). Assume (a). For all $\varepsilon_{l^2} > 0$, there is an $N \geq m$ such that $d_{l^2}(x^{(k)}, x) \leq \varepsilon_{l^2}$ for all $k \geq N$. Let $\varepsilon_{l^1} = \varepsilon_{l^2}/\sqrt{n} > 0$. Then by inequality

(1.1), for all $k \geq N$

$$d_{l^1}(x^{(k)}, x) \leq \sqrt{n} d_{l^2}(x^{(k)}, x) \leq \sqrt{n} \varepsilon_{l^2} = \varepsilon_{l^1}.$$

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

(b) $\Rightarrow$ (c). Assume (b). For all $\varepsilon_{l^1} > 0$, there is an $N \geq m$ such that $d_{l^1}(x^{(k)}, x) \leq \varepsilon_{l^1}$ for all $k \geq N$. Let $\varepsilon_{l^\infty} = \varepsilon_{l^1} > 0$. Then by inequality (1.2), for all $k \geq N$

$$d_{l^\infty}(x^{(k)}, x) \leq d_{l^1}(x^{(k)}, x) \leq \varepsilon_{l^1} = \varepsilon_{l^\infty}.$$

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

(c) $\Rightarrow$ (d). Assume (c). For all $\varepsilon_{l^\infty} > 0$, there is an $N \geq m$ such that $d_{l^\infty}(x^{(k)}, x) \leq \varepsilon_{l^\infty}$ for all $k \geq N$. Let $\varepsilon = \varepsilon_{l^\infty} > 0$. Then by definition, for all $k \geq N$ and all j

$$|x_j^{(k)} - x_j| \leq d_{l^\infty}(x^{(k)}, x) = \sup\{|x_j^{(k)} - x_j : 1 \leq j \leq n|\} \leq \varepsilon_{l^\infty} = \varepsilon.$$

Thus for every $1 \leq j \leq n$, the sequence $(x^{(k)})_{k=m}^{\infty}$ converges to x_j.

(d) $\Rightarrow$ (a). Assume (d). For all $1 \leq j \leq n$ and every $\varepsilon > 0$, there is an $N \geq m$ such that $|x_j^{(k)} - x_j| \leq \varepsilon$ for all $k \geq N$. Let $\varepsilon_{l^2} = \sqrt{n}\varepsilon > 0$. Then we have

$$d(x^{(k)}, x) = \left(\sum_{j=1}^{n} (x_j^{(k)} - x_j)^2 \right)^{1/2} \leq \sqrt{n}\varepsilon = \varepsilon_{l^2}.$$

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

Exercise 1.1.13. *Prove Proposition 1.1.19.*

Proof. Suppose there exists an $N \geq m$ such that $x^{(n)} = x$ for all $n \geq N$. Then for all $\varepsilon > 0$, by definition, we have

$$d_{\text{disc}}(x^{(n)}, x) = 0 \leq \varepsilon,$$

for all $n \geq N$. This means that $(x^{(n)})_{n=m}^{\infty}$ converges to x with respect to the discrete metric d_{disc}.

Conversely, we suppose that $(x^{(n)})_{n=m}^{\infty}$ converges to x with respect to the discrete metric d_{disc}. Then for every $\varepsilon > 0$, there exists an $N \geq m$ such that $d_{\text{disc}}(x^{(k)}, x) \leq \varepsilon$. If $x^{(n)} \neq x$, then $d_{\text{disc}}(x^{(k)}, x) = 1$, for $\varepsilon = 1/2$, we have $d_{\text{disc}}(x^{(k)}, x) > \varepsilon$, a contradiction. Thus $x^{(k)} = x$. $\quad\square$

Exercise 1.1.14. *Prove Proposition 1.1.20. (Hint: modify the proof of Proposition 6.1.7.)*

Proof. Suppose for sake of contradiction that $(a^{(n)})_{n=m}^{\infty}$ was converging to both x and x' with respect to d. Let $\varepsilon = d(x, x')/3$; note that ε is positive since $x \neq x'$. Since $(a^{(n)})_{n=m}^{\infty}$ converges to x, there is an $N \geq m$ such that $d(x^{(n)}, x) \leq \varepsilon$ for all $n \geq N$. Similarly, there is an $M \geq m$ such that $d(x^{(n)}, x') \leq \varepsilon$ for all $m \geq M$. In particular, if we set $n := \max(N, M)$, then we have $d(x^{(n)}, x) \leq \varepsilon$ and $d(x^{(n)}, x') \leq \varepsilon$, hence by the triangle inequality $d(x, x') \leq 2|x - x'|/3$. But then we have $|x - x'| \leq 2|x - x'|/3$, which contradicts the fact that $|x - x'| > 0$. Thus it is not possible to converge to both x and x'. $\square$

Exercise 1.1.15. *Let*
$$X := \left\{ (a_n)_{n=0}^{\infty} : \sum_{n=0}^{\infty} |a_n| < \infty \right\}$$
be the space of absolutely convergent sequences. Define the l^1 and l^{∞} metrics on this space by
$$d_{l^1}((a_n)_{n=0}^{\infty}, (b_n)_{n=0}^{\infty}) := \sum_{n=0}^{\infty} |a_n - b_n|;$$
$$d_{l^{\infty}}((a_n)_{n=0}^{\infty}, (b_n)_{n=0}^{\infty}) := \sup_{n \in \mathbf{N}} |a_n - b_n|.$$
Show that these are both metrics on X, but show that there exist sequences $x^{(1)}, x^{(2)}, \cdots$ of elements of X (i.e., sequences of sequences) which are convergent with respect to the $d_{l^{\infty}}$ metric but not with respect to the d_{l^1} metric. Conversely, show that any sequence which converges in the d_{l^1} metric automatically converges in the $d_{l^{\infty}}$ metric.

Proof. First of all we verify that (X, d_{l^1}) is a metric space. Let $\sum_{n=0}^{N} |a_n|$ be a partial sum of $\sum_{n=0}^{\infty} |a_n|$, and $\lim_{N \to \infty} \sum_{n=0}^{N} |a_n| = A$. This is similar to set $(b_n)_{n=0}^{\infty}, (c_n)_{n=0}^{\infty}, \cdots$ and where they converge to some $B, C, \cdots$.

(a) For all $(a_n)_{n=0}^{\infty} \in X$, we have
$$d_{l^1}((a_n)_{n=0}^{\infty}, (a_n)_{n=0}^{\infty}) = \lim_{N \to \infty} \sum_{n=0}^{N} |a_n - a_n| = \lim_{N \to \infty} 0 = 0.$$

(b) For any distinct $(a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty \in X$, we have

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

$$\geq \left| \sum_{n=0}^\infty |a_n| - \sum_{n=0}^\infty |b_n| \right| = |A - B| > 0.$$

(c) For any $(a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty \in X$, we have

$$\sum_{n=0}^N |a_n - b_n| = \sum_{n=0}^N |b_n - a_n|.$$

Then

$$\lim_{N \to \infty} \sum_{n=0}^N |a_n - b_n| = \lim_{N \to \infty} \sum_{n=0}^N |b_n - a_n|.$$

Thus $d_{l^1}((a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty) = d_{l^1}((b_n)_{n=0}^\infty, (a_n)_{n=0}^\infty)$.

(d) For any $(a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty, (c_n)_{n=0}^\infty \in X$, we have

$$\sum_{n=0}^N |a_n - c_n| \leq \sum_{n=0}^N |(a_n - b_n) + (b_n - c_n)|$$

$$= \sum_{n=0}^N |a_n - b_n| + \sum_{n=0}^N |b_n - c_n|.$$

Then

$$\lim_{N \to \infty} \sum_{n=0}^N |a_n - c_n| \leq \lim_{N \to \infty} \sum_{n=0}^N |a_n - b_n| + \lim_{N \to \infty} \sum_{n=0}^N |b_n - c_n|.$$

Thus

$$d_{l^1}((a_n)_{n=0}^\infty, (c_n)_{n=0}^\infty) \leq d_{l^1}((a_n)_{n=0}^\infty, (c_n)_{n=0}^\infty) + d_{l^1}((b_n)_{n=0}^\infty, (c_n)_{n=0}^\infty).$$

Therefore, (X, d_{l^1}) is a metric space.

Then we verify that (X, d_{l^∞}) is a metric space:

(a) For all $(a_n)_{n=0}^\infty$, we have

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

(b) For any distinct $(a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty \in X$, there is an $N \geq 0$ such that

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

(*c*) For any $(a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty \in X$, we have
$$d_{l^\infty}((a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty) = d_{l^\infty}((b_n)_{n=0}^\infty, (a_n)_{n=0}^\infty)$$
for that $|a_n - b_n| = |b_n - a_n|$ for all $n \in \mathbf{N}$ implies $\sup_{n\in\mathbf{N}} |a_n - b_n| = \sup_{n\in\mathbf{N}} |b_n - a_n|$.

(*d*) For any $(a_n)_{n=0}^\infty, (b_n)_{n=0}^\infty, (c_n)_{n=0}^\infty \in X$. Since $|a_n - c_n| \le |a_n - b_n| + |b_n - c_n|$ for all $n \in \mathbf{N}$, we have
$$\sup_{n\in\mathbf{N}} |a_n - c_n| \le \sup_{n\in\mathbf{N}}(|a_n - b_n| + |b_n - c_n|)$$
$$= \sup_{n\in\mathbf{N}} |a_n - b_n| + \sup_{n\in\mathbf{N}} |b_n - c_n|.$$

Thus (X, d_{l^∞}) is a metric space.

Let $(x^{(k)})_{k=1}^\infty$ denote the sequence
$$x^{(k)} = \left(1 + \frac{1}{k}, \frac{1}{2} + \frac{1}{k}, \cdots, \frac{1}{n} + \frac{1}{k}, \cdots\right).$$

Let $(x_n)_{n=1}^\infty$ denote the sequence
$$x_n = \left(1, \frac{1}{2}, \cdots, \frac{1}{n}, \cdots\right).$$

Then
$$\lim_{k\to\infty} d_{l^\infty}((x^{(k)})_{k=1}^\infty, (x_n)_{n=1}^\infty) = \lim_{k\to\infty} \sup_{n\in\mathbf{N}} \left|\frac{1}{n} + \frac{1}{k} - \frac{1}{n}\right| = \lim_{k\to\infty} \frac{1}{|k|} = 0.$$

But
$$\lim_{k\to\infty} d_{l^1}((x^{(k)})_{k=1}^\infty, (x_n)_{n=1}^\infty) = \lim_{k\to\infty} \sum_{k=1}^\infty \left|\frac{1}{n} + \frac{1}{k} - \frac{1}{n}\right| = \lim_{k\to\infty} \sum_{k=1}^\infty \frac{1}{|k|} = \infty.$$

To show the conclusion in the other direction, suppose that $(x^{(k)})_{k=1}^\infty$ converges to x with respect to d_{l^1}.

Exercise 1.1.16. *Let $(x_n)_{n=1}^\infty$ and $(y_n)_{n=1}^\infty$ be two sequences in a metric space (X, d). Suppose that $(x_n)_{n=1}^\infty$ converges to a point $x \in X$, and $(y_n)_{n=1}^\infty$ converges to a point $y \in X$. Show that $\lim_{n\to\infty} d(x_n, y_n) = d(x, y)$. (Hint: use the triangle inequality several times.)*

Proof. Since $(x_n)_{n=1}^\infty$ and $(y_n)_{n=1}^\infty$ converges to x and y respectively, we have
$$\lim_{n\to\infty} d(x_n, x) = \lim_{n\to\infty} d(y_n, y) = 0.$$

Then

$$\lim_{n\to\infty} d(x_n, y_n) \le \lim_{n\to\infty} \left(d(x_n, y) + d(y_n, y)\right)$$
$$\le \lim_{n\to\infty} \left(d(x_n, x) + d(x, y) + d(y_n, y)\right)$$
$$= \lim_{n\to\infty} d(x_n, x) + \lim_{n\to\infty} d(x, y) + \lim_{n\to\infty} d(y_n, y)$$
$$= d(x, y).$$

$\square$

1.2

Some point-set topology of metric spaces

Exercise 1.2.1. *Verify the claims in Example 1.2.8.*

> **Example 1.2.8.** When we give a set X the discrete metric d_{disc}, and E is any subset of X, then every element of E is an interior point of E, every point not contained in E is an exterior point of E, and there are no boundary points; see Exercise 1.2.1.

Proof. Suppose $x \in E$. Let $r > 0$, for every $x \in X$, we have $d_{\text{disc}}(x, x) = 0 < r$. Then $B(x, r) = \{x\} \subseteq E$, x is an interior point of E. While if $x \notin E$, then we have $d_{\text{disc}}(x, x) = 0 < r$, and we can find a $r > 0$ such that $B(x, r) = \{x\} \cap E = \emptyset$, x is an exterior point of E. Since x is an element of E or not, there is no boundary point. $\square$

Exercise 1.2.2. *Prove Proposition 1.2.10. (Hint: for some of the implications one will need the axiom of choice, as in Lemma 8.4.5.)*

> **Proposition 1.2.10.** Let (X, d) be a metric space, let E be a subset of X, and let x_0 be a point in X. Then the following statements are logically equivalent.
>
> (a) x_0 is an adherent point of E.
>
> (b) x_0 is either an interior point or a boundary point of E.

(c) There exists a sequence $(x_n)_{n=1}^{\infty}$ in E which converges to x_0 with respect to the metric d.

Proof. (a) $\Rightarrow$ (b). Suppose x_0 is an adherent point of E. Then for every radius $r > 0$, the ball $B(x_0, r)$ has a non-empty intersection with E, i.e., $B(x_0, r) \cap E \neq \emptyset$. This means that x_0 is not an exterior point of E. If $B(x_0, r) \cap E = B(x_0, r)$, x_0 is an interior point. Otherwise, it is a boundary point.

(b) $\Rightarrow$ (c). Suppose that x_0 is either an interior point or a boundary point of E. If x_0 is an interior point. Let $x_n = x_0$ for all $n \geq 1$, then we have the sequence $(x_n)_{n=1}^{\infty}$ lies in E and converges to x_0 with respect to the metric d. While if x_0 is a boundary point. Let

$$B_n(x_0, r) := \{x \in X : d(x, x_0) < 1/n\}.$$

Since x_0 is a boundary point, $B_n(x_0, r)$ is not empty for every n ($B(x_0, r) \cap E \neq \emptyset$ for all $r > 0$). Let $\varepsilon = 1/n > 0$, there is an $N \geq 1$ such that $d(x_n, x_0) \leq \varepsilon$ for all $n \geq N$. This means $(x_n)_{n=1}^{\infty}$ converges to x_0.

(c) $\Rightarrow$ (a). Suppose there is a sequence $(x_n)_{n=1}^{\infty}$ in E which converges to x_0 with respect to the metric d. Then for every $r > 0$, we have $d(x_0, x_n) < r$, which means $x_n \in B(x_0, r)$. Therefore, $x_n \in B(x_0, r) \cap E$ and such a set is not an empty set. Thus x_0 is an adherent point of E. $\qquad\square$

Exercise 1.2.3. *Prove Proposition 1.2.15. (Hint: you can use earlier parts of the proposition to prove later ones.)*

Proof. Let (X, d) be a metric space.

(a) *Let E be a subset of X. Then E is open if and only if $E = \text{int}(E)$. In other words, E is open if and only if for every $x \in E$, there exists an $r > 0$ such that $B(x, r) \subseteq E$.*

Suppose E is open. By Definition 1.2.12, $\partial E \cap E = \emptyset$, then for every $x \in E$, we have $x \notin \partial E$. By Definition 1.2.5, x is either an interior or an exterior point. For every $x \in E$, there exists a radius $r > 0$ such that $x \in B(x, r)$. Hence we know that x is an interior, i.e., $B(x, r) \subseteq E$.

Hence $E \subseteq \operatorname{int}(E)$. On the other hand, for every $x \in \operatorname{int}(E)$ are necessary in E. Hence $\operatorname{int}(E) \subseteq E$. We conclude that $E = \operatorname{int}(E)$.

Conversely, if $x \in \partial E$. Since $E = \operatorname{int}(E)$, by Definition 2.1.5, we have $x \notin E$. Thus $\partial E \cap E = \emptyset$ and E is open.

(b) *Let E be a subset of X. Then E is closed if and only if E contains all its adherent points. In other words, E is closed if and only if for every convergent sequence $(x_n)_{n=m}^\infty$ in E, the limit $\lim_{n \to \infty} x_n$ of that sequence also lies in E.*

Suppose E is closed. By Definition 1.2.12, we have $\partial E \subseteq E$. Suppose $x \in E$. Then we have either $x \in \operatorname{int}(E)$ or $x \in \partial E$. Hence $E \subseteq \overline{E}$ by Corollary 1.2.12. Other hand, suppose that $x \in \overline{E}$. Then we have either $x \in \operatorname{int}(E)$ or $x \in \partial E$. If $x \in \operatorname{int}(E)$, then $x \in E$. While if $x \in \partial E$. By Proposition 1.2.10, there exists a sequence $(x_n)_{n=m}^\infty$ in E such that $\lim_{n \to \infty} x_n = x$. Since $\partial E \subseteq E$, we have $x \in E$. Hence $\overline{E} \subseteq E$. We conclude that $E = \overline{E}$.

Conversely, if $E = \overline{E}$. By Corollary 1.2.11, $E = \operatorname{int}(E) \cup \partial E$. Thus we have $\partial E \subseteq E$ and E is closed.

(c) *For any $x_0 \in X$ and $r > 0$, then the ball $B(x_0, r)$ is an open set. The set $\{x \in X : d(x, x_0) \le r\}$ is a closed set. (This set is sometimes called the closed ball of radius r centered at x_0.)*

Suppose $x \in B(x_0, r)$. Let $r > 0$, we have $d(x, x_0) < r$. Then there is $\delta > 0$ such that $0 < \delta < r - d(x, x_0)$ (e.g., let $\delta = (r - d(x, x_0))/3$). That is, $B(x, \delta) \subseteq B(x_0, r)$. Thus for every $x \in B(x_0, r)$, there exists $\delta > 0$ such that $B(x, \delta) \subseteq B(x_0, r)$. By Proposition 1.2.15(a), $B(x_0, r)$ is open.

Let

$$C := \{x \in X : d(x, x_0) \le r\}$$

Suppose $x \in C$. Then for every $r > 0$, $C \cap B(x, r)$ is non-empty (it at least contains x). Hence C contains all of its adherent points. By Proposition 1.2.15(b), C is closed.

(d) *Any singleton set $\{x_0\}$, where $x_0 \in X$, is automatically closed.*

For every $x \in X$, if $x \neq x_0$, then for $r = d(x, x_0) > 0$, we can find $0 < \delta < r$ such that $B(x, \delta) \cap \{x_0\} = \emptyset$. Thus for every $x \in X$ where $x \neq x_0$, we have $x \in \text{ext}(\{x_0\})$. Then we of course have $\overline{\{x_0\}} = X \setminus \text{ext}(\{x_0\}) = \{x_0\}$. By Proposition 2.1.15(b), singleton set $\{x_0\}$ is closed.

(e) *If E is a subset of X, then E is open if and only if the complement $X \setminus E := \{x \in X : x \notin E\}$ is closed.*

Lemma. *If E is a subset of X, then the boundary of E equals to the boundary of the complement of E, i.e., $\partial E = \partial(X \setminus E)$.*

Proof of Lemma. First of all we show that $\text{int}(E) = \text{ext}(X \setminus E)$. Suppose $x \in \text{int}(E)$, for $x \in E$, there exists an $r > 0$ such that $B(x, r) \subseteq E$. Then $B(x, r) \cap (X \setminus E) = \emptyset$. Conversely, suppose that $x \in \text{ext}(X \setminus E)$. For $x \in X \setminus E$, there exists an $r > 0$ such that $B(x, r) \cap (X \setminus E) = \emptyset$. Then $B(x, r) \subseteq E$. Thus $\text{int}(E) = \text{ext}(X \setminus E)$. Since $E = X \setminus (X \setminus E)$, we also have $\text{int}(X \setminus E) = \text{ext}(E)$.

Now we prove (e). By Corollary 1.2.11 and Lemma, we have

$$\partial E = (X \setminus \text{ext}(E)) \setminus \text{int}(E)$$
$$= (X \setminus \text{ext}(X \setminus E)) \setminus \text{int}(X \setminus E) = \partial(X \setminus E).$$

$\square$

Suppose E is open. By Proposition 1.2.15(a) and Corollary 1.2.11, $\partial E \subseteq X \setminus E$. This means that $\partial E = \partial(X \setminus E) \subseteq X \setminus E$. By Definition 1.2.12, $X \setminus E$ is closed. Conversely, suppose that $X \setminus E$ is closed. By Definition 1.2.12, we have $\partial(X \setminus E) \subseteq X \setminus E$. This means that $\partial(X \setminus E) = \partial E \subseteq X \setminus E$, which implies $\partial E \cap (X \setminus E) = \emptyset$. By Definition 1.2.12, E is open.

(f) *If $E_1, \cdots, E_n$ are a finite collection of open sets in X, then $E_1 \cap E_2 \cap \cdots \cap E_n$ is also open. If $F_1, \cdots, F_n$ is a finite collection of closed sets in X, then $F_1 \cup F_2 \cup \cdots \cup F_n$ is also closed.*

For the first part of assertion. Let $x \in E_1 \cap E_2 \cap \cdots \cap E_n$. Then $x \in E_i$ for all $1 \leq i \leq n$. By Proposition 1.2.15(a), there is a radius $r_i > 0$ such that $B(x, r_i) \subseteq E_i$. Let $r = \min(r_1, \cdots, r_n)$, then there exists a radius $r > 0$ such that $B(x, r) \subseteq E_1 \cap E_2 \cap \cdots \cap E_n$. Hence $E_1 \cap E_2 \cap \cdots \cap E_n$ is open.

Then for the second part. By Proposition 1.2.15(e), we prove that $X \setminus (F_1 \cup F_2 \cup \cdots \cup F_n)$ is open. Since we have

$$X \setminus (F_1 \cup F_2 \cup \cdots \cup F_n) = (X \setminus F_1) \cap (X \setminus F_2) \cap \cdots \cap (X \setminus F_n).$$

And since $F_1, \cdots, F_n$ is closed, we have $X \setminus F_1, \cdots, X \setminus F_n$ is open. By the first part assertion, $X \setminus (F_1 \cup F_2 \cup \cdots \cup F_n)$ is open.

(g) *If $\{E_\alpha\}_{\alpha \in I}$ is a collection of open sets in X (where the index set I could be finite, countable, or uncountable), then the union $\bigcup_{\alpha \in I} E_\alpha := \{x \in X : x \in E_\alpha \text{ for some } \alpha \in I\}$ is also open. If $\{F_\alpha\}_{\alpha \in I}$ is a collection of closed sets in X, then the intersection $\bigcap_{\alpha \in I} F_\alpha := \{x \in X : x \in F_\alpha \text{ for all } \alpha \in I\}$ is also closed.*

For the first part of assertion. If $x \in \bigcup_{\alpha \in I} E_\alpha$, then there is $\alpha \in I$ such that $x \in E_\alpha$. Since E_α is open, there is a radius $r > 0$ such that $B(x, r) \subseteq E_\alpha$. Then $B(x, r) \subseteq \bigcup_{\alpha \in I} E_\alpha$. By Proposition 1.2.15, $\bigcup_{\alpha \in I} E_\alpha$ is open.

Then for the second part. We verify that $X \setminus \bigcap_{\alpha \in I} F_\alpha$ is open. Since we have

$$X \setminus \bigcap_{\alpha \in I} F_\alpha = \bigcup_{\alpha \in I} (X \setminus F_\alpha),$$

where $X \setminus F_\alpha$ is open for all $\alpha \in I$. By the first part, we have $X \setminus \bigcap_{\alpha \in I} F_\alpha$ is open.

(h) *If E is any subset of X, then $\mathrm{int}(E)$ is the largest open set which is contained in E; in other words, $\mathrm{int}(E)$ is open, and given any other open set $V \subseteq E$, we have $V \subseteq \mathrm{int}(E)$. Similarly $\overline{E}$ is the smallest closed set which contains E; in other words, $\overline{E}$ is closed, and given any other closed set $K \supset E$, $K \supset \overline{E}$.*

For the first part of assertion. If $x \in V$ and V is an open set, there is a radius $r > 0$ such that $B(x, r) \subseteq V$. So that $B(x, r) \subseteq E$. Which means x is an interior point of E. Thus $x \in \text{int}(E)$ and $V \subseteq \text{int}(E)$.

Then for the second part. We prove that given any open set $X \setminus K \subseteq X \setminus E$, we have $X \setminus K \subseteq X \setminus \overline{E}$ is open. Notice that $X \setminus K$ and $X \setminus \overline{E}$ is closed by Proposition 1.2.15(e). By Lemma, $\text{ext}(E) = \text{int}(X \setminus E)$, and by Corollary, we have $X \setminus \overline{E} = \text{int}(X \setminus E)$. Thus we have $\text{int}(X \setminus E) = X \setminus \overline{E}$ is open, given any other open set $X \setminus K \subseteq X \setminus E$, we have $X \setminus E \subseteq X \setminus \overline{E} = \text{int}(X \setminus \overline{E})$ is open which is hold for the first part of assertion. $\qquad \square$

Exercise 1.2.4. *Let (X, d) be a metric space, x_0 be a point in X, and $r > 0$. Let B be the open ball $B := B(x_0, r) = \{x \in X : d(x, x_0) < r\}$, and let C be the closed ball $C := \{x \in X : d(x, x_0) \leq r\}$.*

(a) Show that $\overline{B} \subseteq C$.

(b) Give an example of a metric space (X, d), a point x_0, and a radius $r > 0$ such that $\overline{B}$ is not equal to C.

Proof. For the part (a). Suppose $x \in \overline{B}$. Then x is an adherent point of B, and for every radius $r > 0$, we have $B(x, r) \cap B \neq \emptyset$. Then $x \in B$ and $d(x, x_0) < r$. This implies that $d(x, x_0) \leq r$. Hence $x \in C$ and $\overline{B} \subseteq C$.

For the part (b). Consider discrete metric. Let $r = 1$, then $B = B(x_0, 1) = \{x_0\}$. And we have $\overline{B} = \{x_0\}$ by Proposition 1.2.15(d). But $C = \{x \in X : d(x, x_0) \leq 1\} = X$. Thus $\overline{B} \neq C$ and $\overline{B} \subsetneq C$. $\qquad \square$

1.3

Relative topology

Exercise 1.3.1. *Prove Proposition 1.3.4(b).*

Proposition 1.3.4. Let (X, d) be a metric space, let Y be a subset of X, and let E be a subset of Y.

(a) E is relatively open with respect to Y if and only if $E = V \cap Y$ for some set $V \subseteq X$ which is open in X.

(b) E is relatively closed with respect to Y if and only if $E = K \cap Y$ for some set $K \subseteq X$ which is closed in X.

Proof. First suppose that E is relatively closed with respect to Y. Then $Y \setminus E$ is open in the metric $(Y, d|_{Y \times Y})$ by Proposition 1.2.15(e). Thus for every $x \in Y \setminus E$, there exists a radius $r_x > 0$ such that the ball $B_{(Y, d|_{Y \times Y})}(x, r_x)$ is contained in $Y \setminus E$.

Now consider the set

$$K := X \setminus \left(\bigcup_{x \in Y \setminus E} B_{(X, d)}(x, r_x) \right).$$

This is a subset of X. By Proposition 1.2.15(c), (g) and (e), K is closed. Now we prove that $E = K \cap Y$. Certainly any point x in E lies in $K \cap Y$, since it lies in Y, and $x \notin Y \setminus E$ implies $x \notin B_{(X, d)}(x, r_x)$, and hence in K. Now suppose that y is a point in $K \cap Y$. Then $y \in K$, which implies that for every $x \in Y \setminus E$ such that $y \notin B_{(X, d)}(x, r_x)$. But since y is also in Y, this implies that $y \notin B_{(Y, d|_{Y \times Y})}(x, r_x)$. But by definition of r_x, this means that $y \notin Y \setminus E$ and $y \in E$, as desired. Thus we have found an open set K for which $E = V \cap Y$ as desired.

Now we do the converse. Suppose that $E = K \cap Y$ for some closed set K; we have to show that $Y \setminus E$ is relatively open with respect to Y. Let x be any point in $Y \setminus E$; we have to show that x is an interior point of $Y \setminus E$ in the metric space $(Y, d|_{Y \times Y})$. Since $x \in Y \setminus E$, we know $x \in X \setminus K$. Since $Y \setminus K$ is open in Y, we know that there is a radius $r > 0$ such that $B_{(Y, d|_{Y \times Y})}(x, r)$ is contained in $Y \setminus K$. Since $E = K \cap Y$, this means that $Y \setminus E = Y \setminus K$, so $B_{(Y, d|_{Y \times Y})}(x, r)$ is contained in $Y \setminus E$. Thus x is an interior point of $Y \setminus E$ in the metric space $(Y, d|_{Y \times Y})$, as desired. $\square$

Cauchy sequences and complete metric spaces

Exercise 1.4.1. *Prove Lemma 1.4.3. (Hint: review your proof of Proposition 6.6.5.)*

> **Lemma 1.4.3.** Let $(x^{(n)})_{n=m}^{\infty}$ be a sequence in (X, d) which converges to some limit x_0. Then every subsequence $(x^{(n_j)})_{j=1}^{\infty}$ of that sequence also converges to x_0.

Proof. Suppose $(x^{(n)})_{n=m}^{\infty}$ converges to x_0. For every $\varepsilon > 0$, there exists an $N \geq m$ such that $d(x^{(n)}, x_0) \leq \varepsilon$ for all $n \geq N$. By Definition 1.4.1, n_j is increasing, we can find an $J \geq 1$ such that $d(x^{(n_j)}, x_0) \leq \varepsilon$ for all $j \geq J$ such that $n_J \geq N$. Hence $(x^{(n_j)})_{j=1}^{\infty}$ converges x_0. $\square$

Exercise 1.4.2. *Prove Proposition 1.4.5. (Hint: review your proof of Proposition 6.6.6.)*

> **Proposition 1.4.5.** Let $(x^{(n)})_{n=m}^{\infty}$ be a sequence of points in a metric space (X, d), and let $L \in X$. Then the following are equivalent:
>
> - L is a limit point of $(x^{(n)})_{n=m}^{\infty}$.
>
> - There exists a subsequence $(x^{(n_j)})_{j=1}^{\infty}$ of the original sequence $(x^{(n)})_{n=m}^{\infty}$ which converges to L.

Proof. Let $n_0 = m$. Define n_j recursively, for every $j \geq 1$,

$$n_j := \min \left\{ n > n_{j-1} : d(x^{(n)}, L) \leq 1/j \right\}.$$

Since L is a limit point of $(x^{(n)})_{n=m}^{\infty}$, by Definition 1.4.4, for every $\varepsilon = 1/j > 0$ and every $N \geq n_0 = m$, there exists an $n \geq N$ such that $d(x^{(n)}, L) \leq 1/j$. Thus $\{n > n_{j-1} : d(x^{(n)}, L) \leq 1/j\}$ is non-empty. By well ordering principle, n_j exists and is unique. Let $x^{(n_j)}$ be a subsequence of $x^{(n)}$. By definition, there exists $N \geq 0$ such that $d(x^{(n_j)}, L) \leq 1/j$ for all $j \geq J$ such that $n_J \geq N$. Hence $(x^{(n_j)})_{j=1}^{\infty}$ converges to L.

For the converse, let $(x^{(n_j)})_{j=1}^\infty$ be a subsequence of $(x^{(n)})_{n=m}^\infty$ which converges to L. By Definition 1.4.1, we can let $n_j = j$ for all $j \geq m$. Since j is a dummy variable, we replace it by n, by Definition 1.4.4, L is a limit point of $(x^n)_{n=m}^\infty$. $\qquad\square$

Exercise 1.4.3. *Prove Lemma 1.4.7. (Hint: review your proof of Proposition 6.1.12.)*

> **Lemma 1.4.7** (Convergent sequences are Cauchy sequences). Let $(x^{(n)})_{n=m}^\infty$ be a sequence in (X, d) which converges to some limit x_0. Then $(x^{(n)})_{n=m}^\infty$ is also a Cauchy sequence.

Proof. Suppose that $(x^{(n)})_{n=m}^\infty$ converges to x_0 in (X, d). By Definition 1.1.14, for every $\varepsilon > 0$, there exists an $N' \geq m$ and $N'' \geq m$ such that $d(x^{(j)}, x_0) \leq \varepsilon$ and $d(x^{(k)}, x_0) \leq \varepsilon$ for all $j \geq N'$ and $j \geq N''$, respectively. Let $N = \max(N', N'')$, we have $N \geq m$ such that $d(x^{(j)}, x^{(k)}) \leq \varepsilon$ for all $j, k \geq N$. By Definition 1.4.6, $(x^{(n)})_{n=m}^\infty$ is a Cauchy sequence. $\qquad\square$

Exercise 1.4.4. *Prove Lemma 1.4.9.*

> **Lemma 1.4.9.** Let $(x^{(n)})_{n=m}^\infty$ be a Cauchy sequence in (X, d). Suppose that there is some subsequence $(x^{(n_j)})_{j=1}^\infty$ of this sequence which converges to a limit x_0 in X. Then the original sequence $(x^{(n)})_{n=m}^\infty$ also converges to x_0.

Proof. Let $\varepsilon > 0$. By Definition 1.4.6, there exists an $N \geq m$ such that $d(x^{(n)}, x^{(j)}) < \varepsilon$ for all $j, n \geq N$. By Definition 1.4.1, there exists an $J \geq m$ such that $d(x^{(n_j)}, x_0) \leq \varepsilon$ for all $j \geq J$. Since $n_j \geq j$, we have $d(x^{(n)}, x^{(n_j)}) < \varepsilon$ for all $n_j, k \geq N$. Thus there exists an $N \geq m$ such that $d(x^{(n)}, x_0) \leq d(x^{(n_j)}, x^{(n)}) + d(x^{(n_j)}, x_0) \leq 2\varepsilon$ for all $n \geq N$. Thus by definition, $(x^{(n)})_{n=m}^\infty$ also converges to x_0. $\qquad\square$

Exercise 1.4.5. *Let $(x^{(n)})_{n=m}^\infty$ be a sequence of points in a metric space (X, d), and let $L \in X$. Show that if L is a limit point of the sequence $(x^{(n)})_{n=m}^\infty$, then L is an adherent point of the set $\{x^{(n)} : n \geq m\}$. Is the*

converse true?

Proof. Since L is a limit point of the sequence $(x^{(n)})_{n=m}^{\infty}$, by Definition 1.4.4, for every $N \geq m$ and $\varepsilon > 0$ there exists an $n \geq N$ such that $d(x^{(n)}, L) \leq \varepsilon$. Define

$$B(L, r) := \{x^{(n)} \in X : d(x^{(n)}, L) < r\}.$$

Let $\varepsilon = r/2$. Then there exists an $n \geq N \geq m$ such that $d(x^{(n)}, L) < r$, and $x^{(n)} \in B(L, r) \cap \{x^{(n)} : n \geq m\}$. By definition, L is an adherent point of $\{x^{(n)} : n \geq m\}$. □

Exercise 1.4.6. *Show that every Cauchy sequence can have at most one limit point.*

Proof. Suppose for sake of contradiction that L and L', where $L \neq L'$, are limit points of $(x^{(n)})_{n=m}^{\infty}$. By Proposition 1.4.5, there exists subsequences $(x^{(n_i)})_{n=m}^{\infty}$ and $(x^{(n_j)})_{n=m}^{\infty}$ converge to L and L' respectively. By Lemma 1.4.9, $(x^{(n)})_{n=m}^{\infty}$ converges to L and L' at the time. But it is impossible from Proposition 1.1.20, a contradiction. □

Exercise 1.4.7. *Prove Proposition 1.4.12.*

> **Proposition 1.4.12.**
>
> (a) Let (X, d) be a metric space, and let $(Y, d|_{Y \times Y})$ be a subspace of (X, d). If $(Y, d|_{Y \times Y})$ is complete, then Y must be closed in X.
>
> (b) Conversely, suppose that (X, d) is a complete metric space, and Y is a closed subset of X. Then the subspace $(Y, d|_{Y \times Y})$ is also complete.

Proof. (a) By Definition 1.4.10, every Cauchy sequence in $(Y, d|_{Y \times Y})$ takes limit in Y. By Lemma 1.4.7, every convergent sequence in Y is Cauchy sequence. By Proposition 1.2.15(b), Y is closed in X.

   (b) By Definition 1.4.10, for every Cauchy sequence in (X, d) in (X, d) takes limit in X. Since $Y \subseteq X$ is closed. By Proposition 1.2.15(b), every sequence in Y takes limit in Y. By Lemma 1.4.7, every Cauchy sequence in $(Y, d|_{Y \times Y})$ takes limit in Y. Thus $(Y, d|_{Y \times Y})$ is also complete. □

Exercise 1.4.8. *The following construction generalizes the construction of the reals from the rationals in Chapter 5, allowing one to view any metric space as a subspace of a complete metric space. In what follows we let (X, d) be a metric space.*

(a) *Given any Cauchy sequence $(x_n)_{n=1}^{\infty}$ in X, we introduce the* formal limit $\mathrm{LIM}_{n\to\infty} x_n$. *We say that two formal limits $\mathrm{LIM}_{n\to\infty} x_n$ and $\mathrm{LIM}_{n\to\infty} y_n$ are equal if $\lim_{n\to\infty} d(x_n, y_n)$ is equal to zero. Show that this equality relation obeys the reflexive, symmetry, and transitive axioms.*

(b) *Let $\overline{X}$ be the space of all formal limits of Cauchy sequences in X, with the above equality relation. Define a metric $d_{\overline{X}} : \overline{X} \times \overline{X} \to \mathbf{R}^+$ by setting*

$$d_{\overline{X}}(\mathrm{LIM}_{n\to\infty} x_n, \mathrm{LIM}_{n\to\infty} y_n) := \lim_{n\to\infty} d(x_n, y_n).$$

Show that this function is well-defined (this means not only that the limit $\lim_{n\to\infty} d(x_n, y_n)$ exists, but also that the axiom of substitution is obeyed; cf. Lemma 5.3.7), and gives $\overline{X}$ the structure of a metric space.

(c) *Show that the metric space $(\overline{X}, d_{\overline{X}})$ is complete.*

(d) *We identify an element $x \in X$ with the corresponding formal limit $\mathrm{LIM}_{n\to\infty} x$ in $\overline{X}$; show that this is legitimate by verifying that $x = y \iff \mathrm{LIM}_{n\to\infty} x = \mathrm{LIM}_{n\to\infty} y$. With this identification, show that $d(x, y) = d_{\overline{X}}(x, y)$, and thus (X, d) can now be thought of as a subspace of $(\overline{X}, d_{\overline{X}})$.*

(e) *Show that the closure of X in $\overline{X}$ is $\overline{X}$ (which explains the choice of notation $\overline{X}$).*

(f) *Show that the formal limit agrees with the actual limit, thus if $(x_n)_{n=1}^{\infty}$ is any Cauchy sequence in X, then we have $\lim_{n\to\infty} x_n = \mathrm{LIM}_{n\to\infty} x_n$ in $\overline{X}$.*

Proof. Let X, d be a metric space. Given any Cauchy sequences $(x_n)_{n=1}^{\infty}$, $(y_n)_{n=1}^{\infty}$ and $(z_n)_{n=1}^{\infty}$ in X.

(a) *Reflexive:* Since $d(x_n, x_n) = 0$, we have $\lim_{n \to \infty} d(x_n, x_n) = 0$. Thus $\text{LIM}_{n \to \infty} x_n = \text{LIM}_{n \to \infty} x_n$.

Symmetry: If $\text{LIM}_{n \to \infty} x_n = \text{LIM}_{n \to \infty} y_n$, we have $\lim_{n \to \infty} d(x_n, y_n) = 0$. Since $d(x_n, y_n) = d(y_n, x_n)$, we have
$$\lim_{n \to \infty} d(x_n, y_n) = \lim_{n \to \infty} d(y_n, x_n) = 0.$$
Thus $\text{LIM}_{n \to \infty} y_n = \text{LIM}_{n \to \infty} x_n$.

Transitive: Suppose that $\text{LIM}_{n \to \infty} x_n = \text{LIM}_{n \to \infty} y_n$ and $\text{LIM}_{n \to \infty} y_n = \text{LIM}_{n \to \infty} z_n$. This means $\lim_{n \to \infty} d(x_n, y_n) = 0$ and $\lim_{n \to \infty} d(y_n, z_n) = 0$. From
$$d(x_n, z_n) \leq d(x_n, y_n) + d(y_n, z_n),$$
we have
$$\lim_{n \to \infty} d(x_n, z_n) \leq \lim_{n \to \infty} d(x_n, y_n) + \lim_{n \to \infty} d(y_n, z_n).$$
Because $d(x_n, z_n) \geq 0$ implies that $\lim_{n \to \infty} d(x_n, z_n) \geq 0$. Thus we have $\lim_{n \to \infty} d(x_n, z_n) = 0$, and $\text{LIM}_{n \to \infty} x_n = \text{LIM}_{n \to \infty} z_n$.

(b) We show that $d_{\overline{X}}$ is well-defined. Let $(x'_n)_{n=1}^{\infty}$ and $(y'_n)_{n=1}^{\infty}$ be Cauchy sequences in X such that $\text{LIM}_{n \to \infty} x_n = \text{LIM}_{n \to \infty} x'_n$ and $\text{LIM}_{n \to \infty} y_n = \text{LIM}_{n \to \infty} y'_n$. Then we have
$$\lim_{n \to \infty} d(x_n, x'_n) = \lim_{n \to \infty} d(y_n, y'_n) = 0.$$
Since
$$d(x_n, y_n) \leq d(x_n, x'_n) + d(x'_n, y'_n) + d(y_n, y'_n),$$
$$d(x'_n, y'_n) \leq d(x'_n, y_n) + d(y_n, x_n) + d(x_n, y'_n).$$
Thus, for every $\varepsilon > 0$, there exists an $N \geq 1$ such that
$$|d(x_n, y_n) - d(x'_n, y'_n)| \leq d(x_n, x'_n) + d(y_n, y'_n) \leq \varepsilon$$
for all $n \geq N$. This shows that $\lim_{n \to \infty} d(x_n, y_n) = \lim_{n \to \infty} d(x'_n, y'_n)$, so that $d_{\overline{X}}(\text{LIM}_{n \to \infty} x_n, \text{LIM}_{n \to \infty} y_n) = d_{\overline{X}}(\text{LIM}_{n \to \infty} x'_n, \text{LIM}_{n \to \infty} y'_n)$.

Now we show that $d_{\overline{X}}$ is a metric on $\overline{X}$, in words, verify Definition 1.1.2:

1) we have $d_{\overline{X}}(\text{LIM}_{n \to \infty} x_n, \text{LIM}_{n \to \infty} x_n) = \lim_{n \to \infty} d(x_n, x_n) = 0$ for that $\text{LIM}_{n \to \infty} x_n = \text{LIM}_{n \to \infty} x_n$.

2) Since $d(x_n, y_n) > 0$, $\lim_{n \to \infty} d(x_n, y_n) > 0$. Hence

$$d_{\overline{X}}(\text{LIM}_{n \to \infty} x_n, \text{LIM}_{n \to \infty} y_n) > 0.$$

3) Since $d(x_n, y_n) = d(y_n, x_n)$, $\lim_{n \to \infty} d(x_n, y_n) = \lim_{n \to \infty} d(y_n, x_n)$. Thus

$$d_{\overline{X}}(\text{LIM}_{n \to \infty} x_n, \text{LIM}_{n \to \infty} y_n) = d_{\overline{X}}(\text{LIM}_{n \to \infty} y_n, \text{LIM}_{n \to \infty} x_n).$$

4) Since $d(x_n, z_n) \leq d(x_n, y_n) + d(y_n, z_n)$, we have

$$\lim_{n \to \infty} d(x_n, z_n) \leq \lim_{n \to \infty} d(x_n, y_n) + \lim_{n \to \infty} d(y_n, z_n).$$

Thus

$$d_{\overline{X}}(\text{LIM}_{n \to \infty} x_n, \text{LIM}_{n \to \infty} z_n) \leq d_{\overline{X}}(\text{LIM}_{n \to \infty} x_n, \text{LIM}_{n \to \infty} y_n)$$
$$+ d_{\overline{X}}(\text{LIM}_{n \to \infty} y_n, \text{LIM}_{n \to \infty} z_n).$$

(*c*) Suppose $(\text{LIM}_{n \to \infty} x_n^{(k)})_{k=1}^{\infty}$ is a Cauchy sequence in $\overline{X}$, we need to show that $(\text{LIM}_{n \to \infty} x_n^{(k)})_{k=1}^{\infty}$ converges in $(\overline{X}, d_{\overline{X}})$. Choose Cauchy sequences $(x_n^{(k)})_{n=1}^{\infty}$ in X, we can see that it also lies in $\overline{X}$. By Definition 1.4.6, for each $k \geq 1$, there exists an $N_k \geq 1$ such that $d(x_n^{(k)}, x_j^{(k)}) < 1/k$ for all $n, j \geq N_k$, so that $\lim_{n \to \infty} d(x_n^{(k)}, x_{N_k}^{(k)}) \leq 1/k$. (In particular, take $x_j^{(k)} = x_{N_k}^{(k)}$, limit exists for the sequence is a Cauchy sequence.) For every k, l, n and assuming $k < l$, we have

$$d(x_{N_k}^{(k)}, x_{N_k}^{(l)}) \leq d(x_{N_k}^{(k)}, x_n^{(k)}) + d(x_n^{(k)}, x_n^{(l)}) + d(x_n^{(l)}, x_{N_k}^{(l)})$$
$$< \frac{1}{k} + d(x_n^{(k)}, x_n^{(l)}) + \frac{1}{l}.$$

Taking the limit as $n \to \infty$, we have

$$d(x_{N_k}^{(k)}, x_{N_k}^{(l)}) \leq \frac{1}{k} + \frac{1}{l} + d_{\overline{X}}(\text{LIM}_{n \to \infty} x_n^{(k)}, \text{LIM}_{n \to \infty} x_n^{(l)}).$$

Because $(\text{LIM}_{n \to \infty} x_n^{(k)})_{k=1}^{\infty}$ is Cauchy, we have $d(x_{N_k}^{(k)}, x_{N_k}^{(l)}) \leq 1/k + 1/l + \zeta$. To prove $(\overline{X}, d_{\overline{X}})$ is complete, we need to show that $\text{LIM}_{n \to \infty} x_n^{(k)} = \text{LIM}_{n \to \infty} x_n$ when $k \to \infty$ in $\overline{X}$. Since $(x_n^{(k)})_{n=1}^{\infty}$ is Cauchy, for every k, there is an $N_k \geq 1$ such that $\lim_{n \to \infty} d(x_n^{(k)}, x_{N_k}^{(k)}) = 0$. And by definition of $\text{LIM}_{n \to \infty} x_n$, we have $\text{LIM}_{n \to \infty} x_n = \text{LIM}_{n \to \infty} x_n^{(n)}$ when

$n \to \infty$. Thus (by substitution)

$$d_{\overline{X}}(\mathrm{LIM}_{n\to\infty} x_n^{(k)}, \mathrm{LIM}_{n\to\infty} x_n) = d_{\overline{X}}(\mathrm{LIM}_{n\to\infty} x_n^{(k)}, \mathrm{LIM}_{n\to\infty} x_{N_k})$$

$$= \lim_{n\to\infty} d(x_n^{(k)}, x_{N_k}^{(n)})$$

$$\leq \lim_{n\to\infty} (d(x_n^{(k)}, x_{N_k}^{(k)}) + d(x_{N_k}^{(k)}, x_{N_k}^{(n)}))$$

$$\leq \lim_{n\to\infty} (d(x_n^{(k)}, x_{N_k}^{(k)}) + d(x_{N_k}^{(k)}, x_{N_k}^{(n)}))$$

$$\leq \frac{1}{2k} + \frac{1}{l} + \zeta.$$

Thus for sufficiently large N_k, we have $d_{\overline{X}}(\mathrm{LIM}_{n\to\infty} x_n^{(k)}, \mathrm{LIM}_{n\to\infty} x_n) \leq \varepsilon$ for all $\varepsilon > 0$, as desired. $\qquad\square$

(d) Suppose that $x = y$, this implies that $\lim_{n\to\infty} d(x, y) = 0$; by definition, $d_{\overline{X}}(\mathrm{LIM}_{n\to\infty} x, \mathrm{LIM}_{n\to\infty} y)$. Thus $\mathrm{LIM}_{n\to\infty} x = \mathrm{LIM}_{n\to\infty} y$. Conversely, suppose that $\mathrm{LIM}_{n\to\infty} x = \mathrm{LIM}_{n\to\infty} y$, then

$$d_{\overline{X}}(\mathrm{LIM}_{n\to\infty} x, \mathrm{LIM}_{n\to\infty} y) = \lim_{n\to\infty} d(x, y) = 0,$$

so that $x = y$, as desired. By this identification, use substitution, we have

$$d_{\overline{X}}(\mathrm{LIM}_{n\to\infty} x, \mathrm{LIM}_{n\to\infty} y) = d_{\overline{X}}(x, y).$$

Thus $d_{\overline{X}}(x, y) = d(x, y)$ and (X, d) can be thought of as a subspace of $\overline{X}, d_{\overline{X}}$.

(e) Let Y to be the closure of X, by Definition 1.2.9, we need to show that for every adherent of X lies in $\overline{X}$.

1.5 Compact metric space

Exercise 1.5.1. *Show that Definitions 9.1.22 and 1.5.3 match when talking about subsets of the real line with the standard metric.*

Definition 9.1.22 (Bounded sets). A subset X of the real line is said to be bounded if we have $X \subseteq [-M, M]$ for some real number $M > 0$.

Definition 1.5.3 (Bounded sets). Let (X, d) be a metric space, and let Y be a subset of X. We say that Y is bounded iff there exists a ball $B(x, r)$ in X which contains Y.

Proof. Let (X, d) be a standard metric space as $(\mathbf{R}, d)$. Let Y be a subset of $\mathbf{R}$. By Definition 1.5.3, Y is bounded implies that there exists a ball $B(x, r) = \{y \in \mathbf{R} : |x - y| < r\}$ for all $y \in Y$. Thus there exists an $M := r + |x|$ such that $Y \subseteq [-M, M]$, as desired. $\qquad\square$

Exercise 1.5.2. *Prove Proposition 1.5.5. (Hint: prove the completeness and boundedness separately. For both claims, use proof by contradiction. You will need the axiom of choice, as in Lemma 8.4.5.)*

Proposition 1.5.5. Let (X, d) be a compact metric space. Then (X, d) is both complete and bounded.

Proof. Since (X, d) is compact, by Definition 1.5.1, there at least one subsequence of any sequence in (X, d) converges in (X, d). Suppose for sake of contradiction that (X, d) is not complete, then by Definition 1.4.10, there exists a Cauchy sequence $(x^{(n)})_{n=m}^{\infty}$ doesn't converge in X. By Lemma 1.4.9, for all subsequence of $(x^{(n)})_{n=m}^{\infty}$ isn't convergent, a contradiction. Thus (X, d) is complete.

Now we suppose that X is not bounded. By Definition 1.5.3, there exists an $x_0 \in X$ but $x_0 \notin B(x, r)$. Let
$$B_n(x, r) := \{x_n \in X : d(x, x_n) < 1/n\}.$$
Then we can take $x_n \in X$ and $d(x, x_n) \geq 1/n$ (i.e., $x_n \notin B_n(x, r)$) for every $n \geq 1$. We can see that $(x_n)_{n=1}^{\infty}$ lies in X which is not convergent. Thus, by Lemma 1.4.3, subsequence of $(x_n)_{n=1}^{\infty}$ also isn't convergent, a contradiction. $\qquad\square$

Exercise 1.5.3. *Prove Theorem 1.5.7. (Hint: use Proposition 1.1.18 and Theorem 9.1.24.)*

Theorem 1.5.7 (Heine-Borel theorem). Let $(\mathbf{R}^n, d)$ be a Euclidean space with either the Euclidean metric, the taxicab metric, or the sup norm metric. Let E be a subset of $\mathbf{R}^n$. Then E is compact if and only if it is closed and bounded.

Proof. By Corollary 1.5.6, E is compact implies that it is closed and bounded. For the converse, suppose that E is closed and bounded. Since $E \subseteq \mathbf{R}^n$, there is $E_i \subseteq \mathbf{R}$ such that $E = \prod_{d=1}^{n} E_d$.[1] Then all of E_d are closed and bounded. Let $(x_i^{(k)})_{i=1}^n$ takes value in E_d such that $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \cdots, x_n^{(k)})$. By Theorem 9.1.24, there is a subsequence of $(x_i^{(k)})_{i=1}^n$ converges in E_d. Let $x^{(k_j)} = (x_{i_1}^{(k)}, x_{i_2}^{(k)}, \cdots, x_{i_n}^{(k)})$. Thus there is a subsequence $(x^{(k_j)})_{j=1}^\infty$ of $(x^{(k)})_{k=m}^\infty$ converges in E and takes limit value in E for that all of its coordinate takes limit in E_d. By Proposition 1.1.18, this conclusion is hold for the Euclidean metric, taxicab metric and sup norm metric. By Definition 1.5.1, E is compact. □

Exercise 1.5.4. *Let $(\mathbf{R}, d)$ be the real line with the standard metric. Give an example of a continuous function $f : \mathbf{R} \to \mathbf{R}$, and an open set $V \subseteq \mathbf{R}$, such that the image $f(V) := \{f(x) : x \in V\}$ of V is* not *open.*

Proof. Consider $f(x) = \sqrt{1 - x^2}$. We can see that $V = (0, 2)$ is open in $\mathbf{R}$ but $f(V) = [0, 1)$ is not open. □

Exercise 1.5.5. *Let $(\mathbf{R}, d)$ be the real line with the standard metric. Give an example of a continuous function $f : \mathbf{R} \to \mathbf{R}$, and a closed set $V \subseteq \mathbf{R}$, such that $f(F)$ is* not *closed.*

Proof. Consider $f(x) = 1/x$. We can see that $F = [0, 1]$ is closed in $\mathbf{R}$ but $f(F) = [1, \infty)$ is not closed. □

Exercise 1.5.6. *Prove Corollary 1.5.9. (Hint: work in the compact metric space $(K_1, d|_{K_1 \times K_1})$, and consider the sets $V_n := K_1 \setminus K_n$, which are open on K_1. Assume for sake of contradiction that $\bigcap_{n=1}^\infty K_n = \emptyset$, and then apply*

---

[1]See *Exercise 3.5.6, Analysis I.*

Theorem 1.5.8.)

> **Corollary 1.5.9.** Let (X, d) be a metric space, and let $K_1, K_2, K_3, \cdots$ be a sequence of non-empty compact subsets of X such that
> $$K_1 \supset K_2 \supset K_3 \supset \cdots.$$
> Then the intersection $\bigcap_{n=1}^{\infty} K_n$ is non-empty.

Proof. Let $(K_1, d|_{K_1 \times K_1})$ be a compact metric space (this works for that $K_1 \subseteq X$ is compact, and by Definition 1.5.1). By Definition 1.5.1 and Proposition 1.5.5, $(K_n, d|_{K_n \times K_n})$, which is subspace of $(K_1, d|_{K_1 \times K_1})$, is complete for all $n \geq 2$, then by Proposition 1.4.12, K_n is closed in X for all $n \geq 2$. Let $V_n := K_1 \setminus K_n$, by Proposition 1.2.15(e), V_n is open for all $n \geq 1$.

Suppose for sake of contradiction that $\bigcap_{n=1}^{\infty} K_n = \emptyset$. Since
$$K_1 \setminus \left(\bigcap_{n=1}^{\infty} K_n \right) = \bigcup_{n=1}^{\infty} (K_1 \setminus K_n) = \bigcup_{n=1}^{\infty} V_n.$$
By hypothesis, $\bigcup_{n=1}^{\infty} V_n = K_1$. By theorem 1.5.8, $K_1 \setminus K_n$ is a compact subset of K_1, and $K_1 \setminus K_n \subseteq \bigcup_{n=1}^{\infty} V_n$, then there is a subcollection such that $K_1 \setminus K_n \subseteq \bigcup_{j=1}^{N} V_{n_j}$. When j takes large enough, we have $K_1 \setminus K_n = K_1$. This means $K_n = \emptyset$, a contradiction. $\square$

Exercise 1.5.7. *Prove Theorem 1.5.10. (Hint: for part (c), you may wish to use (b), and first prove that every singleton set is compact.)*

> **Theorem 1.5.10.** Let (X, d) be a metric space.
>
> (a) If Y is a compact subset of X, and $Z \subseteq Y$, then Z is compact if and only if Z is closed.
>
> (b) If $Y_1, \cdots, Y_n$ are a finite collection of compact subsets of X, then their union $Y_1 \cup \cdots \cup Y_n$ is also compact.
>
> (c) Every finite subset of X (including the empty set) is compact.

Proof. (a) If Z is compact, by Corollary 1.5.6, Z is closed. While if Z is closed. By Proposition 1.2.15(b), the limit of every convergent sequence in Z lies in Z, then by Lemma 1.4.3, their subsequences also lie in Z. By

Definition 1.5.1, Z is compact.

(b) By Theorem 1.5.10(a), all of $Y_1, \cdots, Y_n$ are closed, by Proposition 1.2.15(f), $Y_1 \cup \cdots \cup Y_n$ is also closed. By Theorem 1.5.10(a) again, it is compact.

(c) Let Y be a finite subset of X, and takes $P(n)$ as the property that Y has the cardinality with n. We use induction on n. The base case $n = 0$ is trivial for that $\emptyset$ is closed, so that compact. When $n = 1$, By Proposition 1.2.15(d), singleton also compact. Now we suppose inductively that Y has cardinality with n is compact, and for a singleton $\{x_0\}$ where $x_0 \notin Y$, by Theorem 1.5.7(b) and induction hypothesis, we have $Y \cup \{x_0\}$ is compact. This closed the induction. $\qquad\square$

Exercise 1.5.8. *Let (X, d_{l_1}) be the metric space from Exercise 1.1.15. For each natural number n, let $e^{(n)} = (e_j^{(n)})_{j=0}^\infty$ be the sequence in X such that $e_j^{(n)} := 1$ when $n = j$ and $e_j^{(n)} := 0$ when $n \neq j$. Show that the set $\{e^{(n)} : n \in \mathbf{N}\}$ is a closed and bounded subset of X, but is not compact. (This is despite the fact that (X, d_{l_1}) is even a complete metric space - a fact which we will not prove here. The problem is that not that X is incomplete, but rather that it is "infinite-dimensional", in a sense that we will not discuss here.)*

Proof. Let (X, d_{l^1}) be the metric space. Let

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

be the space of absolutely convergent sequences. Define the l^1 metric on this space by

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

By definition, for every $\varepsilon > 0$ and $j \geq 0$, there exists an $N \geq 0$ such that $e_j = (e_j^{(n)})_{n=0}^\infty$ converges to L_j, i.e.,

$$d_{l^1}(e_j^{(n)}, L_j) = \sum_{n=N}^\infty |e_j^{(n)} - L_j| \leq \varepsilon$$

for all $n \geq N$. We either have j equal to n or not. Thus $(e^{(n)})_{n=m}^{\infty}$ converges to $L := (L_j)_{j=0}^{\infty}$ where $L_j = 1$ when $j = n$ and $L_j = 0$ when $j \neq n$. This is easy to see that L lies in $\{e^{(n)} : n \in \mathbf{N}\}$. Thus $\{e^{(n)} : n \in \mathbf{N}\}$ is closed for that every convergent sequence $(e^{(n)})_{n=m}^{\infty}$ takes limit L in $\{e^{(n)} : n \in \mathbf{N}\}$.

Then we show that $\{e^{(n)} : n \in \mathbf{N}\}$ is bounded. Consider $B(e^{(m)}, 3)$, for every element of $\{e^{(n)} : n \in \mathbf{N}\}$, if $n \neq m$ we have

$$d_{l^1}(e^{(m)}, e^{(n)}) = \sum_{j=0}^{\infty} |e_j^{(m)}, e_j^{(n)}| = |e_m^{(m)}| + |e_n^{(n)}| = 2 < 3.$$

While if $n = m$, we have

$$d_{l^1}(e^{(m)}, e^{(n)}) = \sum_{j=0}^{\infty} |e_j^{(m)}, e_j^{(m)}| = |e_m^{(m)} - e_m^{(m)}| = 0 < 3.$$

Thus $e^{(n)} \in B(e^{(m)}, 3)$ for every $n \in \mathbf{N}$. By definition, $\{e^{(n)} : n \in \mathbf{N}\}$ is bounded.

Since there do not exist an $N \geq 0$ such that for every $n, m \geq N$ and $n \neq m$ such that

$$d_{l^1}(e^n, e^m) = \sum_{j=0}^{\infty} |e_j^{(n)} - e_j^{(m)}| = |e_n^{(n)}| + |e_m^{(m)}| = 2,$$

by Lemmas 1.4.7 and 1.4.9, there is a sequence in $\{e^{(n)} : n \in \mathbf{N}\}$ hasn't convergent subsequence. Thus $\{e^{(n)} : n \in \mathbf{N}\}$ is not compact. $\qquad\square$

Exercise 1.5.9. *Show that a metric space (X, d) is compact if and only if every sequence in X has at least one limit point.*

Proof. Suppose that every sequence in X has at least one limit point, by Proposition 1.4.5, this equivalent to every sequence in X has at least one convergent subsequence. Thus by Definition 1.5.1, (X, d) is compact. For the converse, suppose that (X, d) is compact, then every sequence in X has at least one convergent subsequence, by Proposition 1.4.5, this equivalent to every sequence in X has limit point. $\qquad\square$

Exercise 1.5.10. *A metric space (X, d) is called* totally bounded *if for every $\varepsilon > 0$, there exists a natural number n and a finite number of balls*

$B(x^{(1)}, \varepsilon), \cdots, B(x^{(n)}, \varepsilon)$ *which cover* X *(i.e.,* $X = \bigcup_{i=1}^{n} B(x^{(i)}, \varepsilon))$.[2]

(*a*) *Show that every totally bounded space is bounded.*

(*b*) *Show the following stronger version of Proposition 1.5.5: if* (X, d) *is compact, then complete and totally bounded. (Hint: if* X *is not totally bounded, then there is some* $\varepsilon > 0$ *such that* X *cannot be covered by finitely many* ε-*balls. Then use Exercise 8.5.20 to find an infinite sequence of balls* $B(x^{(n)}, \varepsilon/2)$ *which are disjoint from each other. Use this to then construct a sequence which has no convergent subsequence.)*

(*c*) *Conversely, show that if* X *is complete and totally bounded, then* X *is compact. (Hint: if* $(x^{(n)})_{n=1}^{\infty}$ *is a sequence in* X, *use the total bound-edness hypothesis to recursively construct a sequence of subsequences* $(x^{(n;j)})_{n=1}^{\infty}$ *of* $(x^{(n)})_{n=1}^{\infty}$ *for each positive integer* j, *such that for each* j, *the elements of the sequence* $(x^{(n;j)})_{n=1}^{\infty}$ *are contained in a single ball of radius* $1/j$, *and also that each sequence* $(x^{(n;j+1)})_{n=1}^{\infty}$ *is a subsequence of the previous one* $(x^{(n;j)})_{n=1}^{\infty}$. *Then show that the "diagonal" sequence* $(x^{(n;n)})_{n=1}^{\infty}$ *is a Cauchy sequence, and then use the completeness hypothesis.)*

Proof. (a) Suppose that (X, d) is totally bounded, by definition, for every $\varepsilon > 0$, there exists a positive integer n and a finite number of balls $B(x^{(1)}, \varepsilon), \cdots, B(x^{(n)}, \varepsilon)$ such that $X = \bigcup_{i=1}^{n} B(x^{(i)}, \varepsilon)$. We takes $R = \varepsilon + \max_{1 \le i \le n} d(x^{(i)}, x^{(1)})$. By triangle inequality, we have

$$d(x^{(1)}, x) \le d(x^{(1)}, x^{(i)}) + d(x^{(i)}, x) \le R$$

for all $1 \le i \le n$ and $x \in X$. Thus we have $B(x^{(i)}, \varepsilon) \subseteq B(x^{(1)}, R)$ for all $1 \le i \le n$. Since $X = \bigcup_{i=1}^{n} B(x^{(i)}, \varepsilon)$, X is bounded for that $X \subseteq B(x^{(1)}, R)$.

(b) By Proposition 1.5.5, (X, d) is compact implies complete, we only need to show that it is totally bounded. Suppose for sake of contradiction that (X, d) is not totally bounded. By definition, there is some $\varepsilon > 0$ such

---

[2]In Exercise 1.5.10, n should be a natural number rather than a positive integer (in order to ensure that the empty set is totally bounded). — Errata from Tao.

that there is an $x \in X$ but $x \notin \bigcup_{i=1}^{n} B(x^{(i)}, \varepsilon)$ for all positive integer n. By Exercise 8.5.20 [3], $\bigcup_{i=1}^{n} B(x^{(i)}, \varepsilon) \subseteq 2^{X}$ is a collection of subsets of X, then there is a subset $\bigcup_{i=1}^{n} B(x^{(i)}, \varepsilon/2) \subseteq \bigcup_{i=1}^{n} B(x^{(i)}, \varepsilon)$ such that all of the elements of $\bigcup_{i=1}^{n} B(x^{(i)}, \varepsilon/2)$ are disjoint from each other. For the similar process, we can take subsets $\bigcup_{i=1}^{\infty} B(x^{(i)}, \varepsilon/m)$ for $m \geq 1$. From this, we can see that $(x^{(n)})_{n=1}^{\infty}$ lies in X but is not a Cauchy sequence, by Lemmas 1.4.7 and 1.4.9, this sequence hasn't a convergent subsequence. By definition, (X, d) is not compact, a contradiction.

(c) Suppose that X is complete and totally bounded. Let $(x^{(n)})_{n=1}^{\infty}$ be a sequence in X. Since X is totally bounded, by definition, we have $X = \bigcup_{n=1}^{N} B(x^{(n)}, 1)$. Then must one of $B(x^{(n)}, 1)$ containing the infinitely many terms, which takes the subsequence $(x^{(n;1)})_{n=1}^{\infty}$ of $(x^{(n)})_{n=1}^{\infty}$ to be element. Since X is totally bounded and $B(x^{(n)}, 1) \subseteq X$, $B(x^{(n)}, 1)$ is also totally bounded. Hence we have

$$B(x^{(n)}, 1) = \bigcup_{n=1}^{N} B(x^{(n;1)}, 1/2).$$

Then must one of $B(x^{(n;1)}, 1/2)$ containing infinitely many terms, which takes subsequence $(x^{(n;2)})_{n=1}^{\infty}$ of $(x^{(n;1)})_{n=1}^{\infty}$ to be element. Continuing in this process, we have the subsequence $(x^{(n;j+1)})_{n=1}^{\infty}$ of $(x^{(n;j)})_{n=1}^{\infty}$ lies in $B(x^{(n;j)}, \frac{1}{j+1})$. Consider $(x^{(n;n)})_{n=1}^{\infty}$ which lies in $B(x^{(n;n-1)}, \frac{1}{n})$, for every $x \in B(x^{(n;n-1)}, \frac{1}{n})$, we have

$$d(x^{(k;k)}, x^{(l;l)}) \leq d(x^{(k;k)}, x) + d(x^{(l;l)}, x) \leq \frac{2}{n}$$

for all $k, l \geq n$. Thus $(x^{(n;n)})_{n=1}^{\infty}$ is a Cauchy sequence. Since X complete, the subsequence is convergent. Thus X is compact. $\square$

Exercise 1.5.11. *Let (X, d) have the property that every open cover of X has a finite subcover. Show that X is compact. (Hint: if X is not compact, then by Exercise 1.5.9, there is a sequence $(x^{(n)})_{n=1}^{\infty}$ with no limit points. Then for every $x \in X$ there exists a ball $B(x, \varepsilon)$ containing x which contains at most finitely many elements of this sequence. Now use the hypothesis.)*

---

[3]See *Exercise 8.5.20, Analysis I.*

Proof. Suppose for sake of contradiction that X is not compact. Then by Exercise 1.5.9, there is a sequence $(x^{(n)})_{n=1}^{\infty}$ with no limit points. For every $x \in X$ there exists a ball $B(x, \varepsilon)$ which contains at most finitely many elements of the sequence. By hypothesis, we can choose $\varepsilon_n > 0$ for every n such that $X \subseteq \bigcup_{i=1}^{n} B(x_n, \varepsilon_n)$. This means that the sequence contains finite elements in X, a contradiction. $\qquad\square$

1.6 Appendix

Given a metric space, one can then define various useful topological struc-
tures. One is via the machinery of convergent sequences. There is two
approach to define such almost the same topological structure.

Definition 1 (W. Rudin). Let (X, d) be a metric space, let E be a
subset of X.

(a) A point x_0 is a *limit point* of the set E if for every $r > 0$, the ball
$B(x_0, r)$ has a non-empty intersection with $E \setminus \{x_0\}$.

(b) If $x \in E$ and x is not a limit point of E, then x is called an *isolated
point* of E.

(c) The set of all limit points of E is denoted by E', then the *closure*
of E is the set $\overline{E} = E \cup E'$.

(d) x_0 is an *interior point* of E if there exists a radius $r > 0$ such that
$B(x_0, r) \subset E$.

(e) E is *closed* if every limit point of E is a point of E.

(f) E is *open* if every point of E is an interior point of E.

(g) E is *dense* in X if every point of X is a limit point of E, or a point
of E, or both.

Definition 2 (T. Tao). Let (X, d) be a metric space, let E be a subset
of X.

(a) x_0 is an *adherent point* of E if for every radius $r > 0$, the ball
$B(x_0, r)$ has a non-empty intersection with E. Every adherent point
is either a limit point or an isolated point of E.

(b) The set of all adherent points of E is called the *closure* of E and is
denoted $\overline{E}$.

(c) x_0 is an *interior point* of E if there exists a radius $r > 0$ such that $B(x_0, r) \subset E$.

(d) x_0 is an *exterior point* of E if there exists a radius $r > 0$ such that $B(x_0, r) \cap E = \emptyset$.

(e) x_0 is a *boundary point* of E if it is neither an interior point nor an exterior point of E.

(f) E is *closed* if it contains all of its boundary points, i.e., $\partial E \subset E$. Since an adherent point of E is an interior point or a boundary point of E, we also call a set E is closed if it contains all its adherent points, i.e. if $E = \overline{E}$.

(g) E is *open* if it contains none of its boundary points, i.e., $\partial E \cap E = \emptyset$.

(h) E is *dense* if every point in X is adherent to E, or equivalently if $\overline{E} = X$.

Continuous functions on metric spaces

Continuous functions

Exercise 2.1.1. *Prove Theorem 2.1.4. (Hint: review your proof of Proposition 9.4.7.)*

Theorem 2.1.4 (Continuity preserves convergence). Suppose that (X, d_X) and (Y, d_Y) are metric spaces. Let $f : X \to Y$ be a function, and let $x_0 \in X$ be a point in X. Then the following three statements are logically equivalent:

(*a*) f is continuous at x_0.

(*b*) Whenever $(x^{(n)})_{n=1}^{\infty}$ is a sequence in X which converges to x_0 with respect to the metric d_X, the sequence $(f(x^{(n)}))_{n=1}^{\infty}$ converges to $f(x_0)$ with respect to the metric d_Y.

(*c*) For every open set $V \subset Y$ that contains $f(x_0)$, there exists an open

set $U \subset X$ containing x_0 such that $f(U) \subseteq V$.

Proof. (a) $\Rightarrow$ (b). Suppose that f is continuous at x_0, by Definition 2.1.1, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d_Y(f(x), f(x_0)) < \varepsilon$ whenever $d_X(x, x_0) < \delta$. Since $(x^{(n)})_{n=1}^{\infty}$ converges to x_0 in X, by Definition 1.1.14, for every $\delta > 0$, there exists an $N \geq 1$ such that $d_X(x^{(n)}, x_0) \leq \delta$ for all $n \geq N$. Thus for every $\varepsilon > 0$, there exists an $N \geq 1$ such that $d_Y(f(x^{(n)}), f(x_0)) \leq \varepsilon$ for all $n \geq N$. By Definition 1.1.14 again, $(f(x^{(n)}))_{n=1}^{\infty}$ converges to $f(x_0)$ with respect to the metric d_Y.

(b) $\Rightarrow$ (c). Since V is open in Y and contains $f(x_0)$, by Proposition 1.2.15(a), there exists a $r > 0$ such that $B(f(x_0), r) \subseteq V$. For $\varepsilon = r/2$ and $\delta > 0$, there exists an $N > 1$ such that $d_Y(f(x^{(n)}), f(x_0)) < r/2$ for all $n > N$ whenever $d_X(x^{(n)}, x_0) < \delta$ for all $n > N$. Thus there exists an open set $U = B(x_0, \delta) \subseteq X$ which containing $x^{(n)}$ for such that $n > N$, hence such that $f(U) \subseteq V$.

(c) $\Rightarrow$ (a). For every $\varepsilon > 0$ and $\delta > 0$, $B(f(x_0), \varepsilon)$ and $B(x_0, \delta)$ is an open set in Y and in X, respectively, by Proposition 1.2.15(c). Then for every $\varepsilon > 0$ such that $B(f(x_0), \varepsilon) \subset Y$, there exists a $\delta > 0$ such that $B(x_0, \delta) \subset X$, we have $f(B(x_0, \delta)) \subseteq B(f(x_0), \varepsilon)$. This means, by definition of ball, for every $\varepsilon > 0$ there exists a $\delta > 0$ for which if $d_X(x, x_0) < \delta$, then $d_Y(f(x), f(x_0)) < \varepsilon$. Thus f is continuous at x_0 by Definition 2.1.1. □

Exercise 2.1.2. *Prove Theorem 2.1.5. (Hint: Theorem 2.1.4 already shows that (a) and (b) are equivalent.)*

Theorem 2.1.5. Let (X, d_X) be a metric space, and let (Y, d_Y) be another metric space. Let $f : X \to Y$ be a function. Then the following four statements are equivalent:

(*a*) f is continuous.

(*b*) Whenever $(x^{(n)})_{n=1}^{\infty}$ is a sequence in X which converges to some point $x_0 \in X$ with respect to the metric d_X, the sequence $(f(x^{(n)}))_{n=1}^{\infty}$ converges to $f(x_0)$ with respect to the metric d_Y.

(c) Whenever V is an open set in Y, the set $f^{-1}(V) := \{x \in X : f(x) \in V\}$ is an open set in X.

(d) Whenever F is a closed set in Y, the set $f^{-1}(F) := \{x \in X : f(x) \in F\}$ is a closed set in X.

Proof. (a) $\Leftrightarrow$ (b). This follows immediately from Theorem 2.1.4.

(a) $\Rightarrow$ (c). Suppose that f is continuous. Let V be an open subset of Y. Let x be an element of $f^{-1}(V)$. We need to show that there is a ball $B(x, r)$ contained by $f^{-1}(V)$. Since $f(x) \in V$, by Theorem 2.1.14, there is an open set $B(x, r) \subseteq X$ such that $f(B(x, r)) \subseteq V$ for some $r > 0$. This implies that $B(x, r) \subseteq f^{-1}(V)$, as desired.

(c) $\Rightarrow$ (d). Let $F := Y \setminus V$, we can see that F is a closed set in Y. Then we have $f^{-1}(F) = f^{-1}(Y \setminus V) = \{x \in X : f(x) \in Y \setminus V\}$ and $f^{-1}(F) = f^{-1}(Y) \setminus f^{-1}(V) = X \setminus f^{-1}(V)$.[1] Because $f^{-1}(V)$ is open, hence $f^{-1}(F)$ is closed.

(d) $\Rightarrow$ (a). By Proposition 1.2.15(b), for every closed set F in Y, every sequence $(f(x_n))_{n=m}^{\infty}$ in F has limit lies in F. Then we can choose a sequence $(x_n)_{n=m}^{\infty}$ from $f^{-1}(F)$ such that when $(x_n)_{n=m}^{\infty}$ converges to x_0, $(f(x_n))_{n=m}^{\infty}$ converges to $f(x_0)$. Thus f is continuous. $\square$

Exercise 2.1.3. *Use Theorem 2.1.4 and Theorem 2.1.5 to prove Corollary 2.1.7.*

Corollary 2.1.7 (Continuity preserved by composition). Let (X, d_X), (Y, d_Y), and (Z, d_Z) be metric spaces.

(a) If $f : X \to Y$ is continuous at a point $x_0 \in X$, and $g : Y \to Z$ is continuous at $f(x_0)$, then the composition $g \circ f : X \to Z$, defined by $g \circ f(x) := g(f(x))$, is continuous at x_0.

(b) If $f : X \to Y$ is continuous, and $g : Y \to Z$ is continuous, then $g \circ f : X \to Z$ is also continuous.

---

[1]See *Exercise 3.4.4, Analysis I.*

Proof. (a) Since f is continuous x_0, by Theorem 2.1.4, for every open set $U \subseteq Y$ that contains $f(x_0)$, there exists an open set $U \subseteq X$ containing x_0 such that $f(U) \subseteq V$. Since g is continuous at $f(x_0)$, by Theorem 2.1.4, for every open set $W \subseteq Z$ that contains $g(f(x_0))$, there exists an open set $V \subseteq Y$ containing $f(x_0)$ such that $f(V) \subseteq W$. Thus, by Theorem 2.1.4, for every open set $W \subseteq Z$ that contains $g(f(x_0))$, there exists an open set $U \subseteq X$ containing x_0 such that $g(f(U)) \subseteq W$.

(b) Let W be open in Z. By Theorem 2.1.5, since g is continuous, $g^{-1}(W)$ is open in Y and therefore, since f is continuous, $f^{-1}(g^{-1}(W)) = (g \circ f)^{-1}(W)$ is open in X. Therefore, $g \circ f$ is continuous. □

Exercise 2.1.4. *Give an example of functions $f : \mathbf{R} \to \mathbf{R}$ and $g : \mathbf{R} \to \mathbf{R}$ such that*

(a) *f is not continuous, but g and $g \circ f$ are continuous;*

(b) *g is not continuous, but f and $g \circ f$ are continuous;*

(c) *f and g are not continuous, but $g \circ f$ is continuous.*

Explain briefly why these examples do not contradict Corollary 2.1.7.

Proof.

(a) Consider $f : \mathbf{R} \to \mathbf{R}$ defined as

$$f(x) := \begin{cases} 1 & x > 0, \\ -1 & x \leq 0. \end{cases}$$

and $g : \mathbf{R} \to \mathbf{R}$ defined as $g(y) := y^2$. We can see that $g \circ f$ is continuous.

(b) Consider $f : \mathbf{R} \to \mathbf{R}$ defined as $f(x) := e^x$ and $g : \mathbf{R} \to \mathbf{R}$ defined as $g(y) := 1/y$. We can see that $g \circ f$ is continuous.

(c) Consider $f : \mathbf{R} \to \mathbf{R}$ defined as $f(x) := 1/x$ and $g : \mathbf{R} \to \mathbf{R}$ defined as $g(y) := 1/y$. We can see that $g \circ f$ is continuous. □

Exercise 2.1.5. *Let (X, d) be a metric space, and let $(E, d|_{E \times E})$ be a subspace of (X, d). Let $\iota_{E \to X} : E \to X$ be the inclusion map, defined by setting $\iota_{E \to X}(x) := x$ for all $x \in E$. Show that $\iota_{E \to X}$ is continuous.*

Proof. For every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$d_Y(\iota_{E \to X}(x), \iota_{E \to X}(x_0)) = d_Y(x, x) = 0 < \varepsilon$$

whenever $d_X(x, x_0) < \delta$. Thus by definition, $\iota_{E \to X}$ is continuous. $\square$

Exercise 2.1.6. *Let $f : X \to Y$ be a function from one metric space (X, d_X) to another (Y, d_Y). Let E be a subset of X (which we give the induced metric $d_X|_{E \times E}$), and let $f|_E : E \to Y$ be the restriction of f to E, thus $f|_E(x) := f(x)$ when $x \in E$. If $x_0 \in E$ and f is continuous at x_0, show that $f|_E$ is also continuous at x_0. (Is the converse of this statement true? Explain.) Conclude that if f is continuous, then $f|_E$ is continuous. Thus restriction of the domain of a function does not destroy continuity. (Hint: use Exercise 2.1.5.)*

Proof. Since f is continuous at x_0, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$d_Y(f|_E(x), f|_E(x_0)) = d_Y(f(x), f(x_0)) < \varepsilon$$

whenever $d_X(x, x_0) < \delta$. Because $E \times E \subseteq X \times X$, we have $d_X|_{E \times E}(x, x_0) < \delta$. Therefore, $f|_E$ is continuous at x_0. The converse is not true for that $d_X|_{E \times E}(x, x_0) < \delta$ cannot imply $d_X(x, x_0) < \delta$.

Let (X, d) be a metric space, and let $(E, d_X|_{E \times E})$ be a subspace of (X, d). Let $f : X \to Y$ to be continuous. By Exercise 2.1.5, $\iota_{E \to X} : E \to X$ also is continuous. Thus by Corollary 2.1.7(b), $f_E = f \circ \iota_{E \to X} : E \to Y$ is also continuous. $\square$

Exercise 2.1.7. *Let $f : X \to Y$ be a function from one metric space (X, d_X) to another (Y, d_Y). Suppose that the image $f(X)$ of X is contained in some subset $E \subset Y$ of Y. Let $g : X \to E$ be the function which is the same as f but with the range restricted from Y to E, thus $g(x) = f(x)$ for all $x \in X$. We give E the metric $d_Y|_{E \times E}$ induced from Y. Show that for*

any $x_0 \in X$, that f is continuous at x_0 if and only if g is continuous at x_0. Conclude that f is continuous if and only if g is continuous. (Thus the notion of continuity is not affected if one restricts the range of the function.)

Proof. Suppose that f is continuous at x_0. Since $E \times E \subset Y \times Y$ and $g(x) = f(x)$ for all $x \in X$, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d_Y(f(x), f(x_0)) < \varepsilon$, so that $d_Y|_{E \times E}(g(x), g(x_0)) < \varepsilon$ whenever $d_X(x, x_0) < \delta$. Thus g is continuous at x_0. Conversely, suppose that g is continuous at x_0. Since $f(X) \subseteq E$, for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $d_Y|_{E \times E}(g(x), g(x_0)) < \varepsilon$, so that $d_Y(f(x), f(x_0)) < \varepsilon$ whenever $d_X(x, x_0) < \delta$. Thus f is continuous. $\qquad\square$

2.2 Continuity and product spaces

Exercise 2.2.1. *Prove Lemma 2.2.1. (Hint: use Proposition 1.1.18 and Theorem 2.1.4.)*

> **Lemma 2.2.1.** Let $f : X \to \mathbf{R}$ and $g : X \to \mathbf{R}$ be functions, and let $f \oplus g : X \to \mathbf{R}^2$ be their direct sum. We give $\mathbf{R}^2$ the Euclidean metric.
>
> (a) If $x_0 \in X$, then f and g are both continuous at x_0 if and only if $f \oplus g$ is continuous at x_0.
>
> (b) f and g are both continuous if and only if $f \oplus g$ is continuous.

Proof. (a) Since f and g are continuous at x_0, by Theorem 2.1.4, for all $(x^{(n)})_{n=1}^{\infty}$ in X converges to x_0, $(f(x^{(n)}))_{n=1}^{\infty}$ and $(g(x^{(n)}))_{n=1}^{\infty}$ converges to $f(x_0)$ and $g(x_0)$ respectively. Let $(f(x^{(n)}), g(x^{(n)}))_{n=1}^{\infty}$ be a sequence in $\mathbf{R}^2$. By Proposition 1.1.18 ((d) implies (a)), $(f(x^{(n)}), g(x^{(n)}))_{n=1}^{\infty}$ converges to $(f(x_0), g(x_0))$. Thus $f \oplus g$ is continuous at x_0. While if $f \oplus g$ is continuous at x_0, by Theorem 1.1.4, for every $(x^{(n)})_{n=1}^{\infty}$ in X converges to x_0, $(f(x^{(n)}), g(x^{(n)}))_{n=1}^{\infty}$ converges to $(f(x_0), g(x_0))$. By Proposition 1.1.18 ((a) implies (d)), $(f(x^{(n)}))_{n=1}^{\infty}$ and $(g(x^{(n)}))_{n=1}^{\infty}$ converges to $f(x_0)$ and $g(x_0)$

respectively. Thus f and g are continuous.

(b) is immediately follows from (a). $\qquad\square$

Exercise 2.2.2. *Prove Lemma 2.2.2. (Hint: use Theorem 2.1.5 and limit laws (Theorem 6.1.19).)*

> **Lemma 2.2.2.** The addition function $(x, y) \mapsto x + y$, the subtraction function $(x, y) \mapsto x - y$, the multiplication function $(x, y) \mapsto xy$, the maximum function $(x, y) \mapsto \max(x, y)$, and the minimum function $(x, y) \mapsto \min(x, y)$, are all continuous functions from $\mathbf{R}^2$ to $\mathbf{R}$. The division function $(x, y) \mapsto x/y$ is a continuous function from $\mathbf{R} \times (\mathbf{R} \setminus \{0\}) = \{(x, y) \in \mathbf{R}^2 : y \neq 0\}$ to $\mathbf{R}$. For any real number c, the function $x \mapsto cx$ is a continuous function from $\mathbf{R}$ to $\mathbf{R}$.

Proof. Let $(x^{(n)}, y^{(n)})_{n=1}^{\infty}$ converges to (x_0, y_0) in $\mathbf{R}^2$. Consider addition function $(x, y) \mapsto x + y$. Whenever $(x^{(n)}, y^{(n)})_{n=1}^{\infty}$ converges to (x_0, y_0) in $\mathbf{R}^2$ with respect to the metric $d_{\mathbf{R}^2}$, the sequence $(x_n + y_n)_{n=1}^{\infty}$ converges to $x_0 + y_0$ with respect to the metric $d_{\mathbf{R}}$, by Theorem 6.1.19. By Theorem 2.1.5, $(x, y) \mapsto x + y$ is continuous. A similar argument gives the continuity of $(x, y) \mapsto x - y$, $(x, y) \mapsto xy$, $(x, y) \mapsto \max(x, y)$, $(x, y) \mapsto \min(x, y)$ and $x \mapsto cx$. And by Exercise 2.1.7, $(x, y) \mapsto x/y$ follows immediately. $\qquad\square$

Remark 2.2.2. Since functions above map between $(\mathbf{R}^2, d_{\mathbf{R}^2})$ and $(\mathbf{R}, d_{\mathbf{R}})$, and $(\mathbf{R}, d_{\mathbf{R}})$ equivalent to $(\mathbf{R}, d)$, we can use conclusions from *Analysis I*. Thus Theorem 6.1.19 preserves $d_{\mathbf{R}}(x + y, x_0 + y_0) < \varepsilon$ for all $\varepsilon > 0$, etc.

Exercise 2.2.3. *Show that if $f : X \to \mathbf{R}$ is a continuous function, so is the function $|f| : X \to \mathbf{R}$ defined by $|f|(x) := |f(x)|$.*

Proof. Let f to be a continuous function. Let $g : \mathbf{R} \to \mathbf{R}$ defined by $g(x) := |x|$. We can see that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$||x| - |x_0|| \leq |x - x_0| < \varepsilon$$

whenever $|x - x_0| < \delta = \varepsilon$ for all $x_0 \in \mathbf{R}$. This implies that g is continuous. Thus by Corollary 2.1.7(b), $|f| = g \circ f$ is continuous. $\qquad\square$

Exercise 2.2.4. *Let $\pi_1 : \mathbf{R}^2 \to \mathbf{R}$ and $\pi_2 : \mathbf{R}^2 \to \mathbf{R}$ be the functions $\pi_1(x,y) := x$ and $\pi_2(x,y) := y$ (these two functions are sometimes called the co-ordinate functions on $\mathbf{R}^2$). Show that π_1 and π_2 are continuous. Conclude that if $f : \mathbf{R} \to X$ is any continuous function into a metric space (X,d), then the functions $g_1 : \mathbf{R}^2 \to X$ and $g_2 : \mathbf{R}^2 \to X$ defined by $g_1(x,y) := f(x)$ and $g_2(x,y) := f(y)$ are also continuous.*

Proof. Suppose that $(x^{(n)})_{n=m}^{\infty}$ and $(y^{(n)})_{n=m}^{\infty}$ to be two sequences in $\mathbf{R}$ which converges to x_0 and y_0 respectively. By Proposition 1.1.18, $(x^{(n)}, y^{(n)})_{n=m}^{\infty}$ converges to (x_0, y_0) in $\mathbf{R}^2$ (with respect to the metric d_{l^1}, d_{l^2} and d_{l^∞}). Thus we have the inequality

$$d(\pi_1(x,y), x_0) \leq d((x,y), (x_0, y_0)),$$
$$d(\pi_2(x,y), y_0) \leq d((x,y), (x_0, y_0)).$$

This means that $(\pi_1(x^{(n)}, y^{(n)}))_{n=m}^{\infty}$ converges to x_0 and $(\pi_2(x^{(n)}, y^{(n)}))_{n=m}^{\infty}$ converges to y_0. Thus π_1 and π_2 are continuous. Since $g_i = \pi_i \circ f$, by Corollary 2.1.7, g_1 and g_2 are continuous. $\square$

Exercise 2.2.5. *Let $n, m \geq 0$ be integers. Suppose that for every $0 \leq i \leq n$ and $0 \leq j \leq m$ we have a real number c_{ij}. Form the function $P : \mathbf{R}^2 \to \mathbf{R}$ defined by*

$$P(x,y) := \sum_{i=0}^{n} \sum_{j=0}^{m} c_{ij} x^i y^j.$$

(Such a function is known as a polynomial of two variables; a typical example of such a polynomial is $P(x,y) = x^3 + 2xy^2 - x^2 + 3y + 6$.) Show that P is continuous. (Hint: use Exercise 2.2.4 and Corollary 2.2.3.) Conclude that if $f : X \to \mathbf{R}$ and $g : X \to \mathbf{R}$ are continuous functions, then the function $P(f,g) : X \to \mathbf{R}$ defined by $P(f,g)(x) := P(f(x), g(x))$ is also continuous.

Proof. By Exercise 2.2.4, $\pi_1(x,y) = x$ and $\pi_2(x,y) = y$ are continuous. By Corollary 2.2.3, $\pi_1^i(x,y) = x^i$ and $\pi_2^j(x,y) = y^j$ are continuous, so that $(c_{ij} \pi_1^i \pi_2^j)(x,y) = c_{ij} x^i y^j$ is continuous. By Corollary 2.2.3 again, $P(x,y) := \sum_{i=0}^{n} \sum_{j=0}^{m} c_{ij} x^i y^j$ is continuous. By Lemma 2.2.1, $f \oplus g : X \to \mathbf{R}^2$ is continuous. By Corollary 2.1.7, $P(f,g) = (f \oplus g) \circ P$ is continuous. $\square$

Exercise 2.2.6. *Let $\mathbf{R}^m$ and $\mathbf{R}^n$ be Euclidean spaces. If $f : X \to \mathbf{R}^m$ and $g : X \to \mathbf{R}^n$ are continuous functions, show that $f \oplus g : X \to \mathbf{R}^{m+n}$ is also continuous, where we have identified $\mathbf{R}^m \times \mathbf{R}^n$ with $\mathbf{R}^{m+n}$ in the obvious manner. Is the converse statement true?*

Proof. We can fixed $m = 1$ and use induction on n, by Lemma 2.2.1, we have $f : X \to \mathbf{R}$ and $g : X \to \mathbf{R}^n$ are both continuous if and only if $f \oplus g : X \to \mathbf{R}^{1+n}$ is continuous. Then given n and induct on m, we can conclude that $f : X \to \mathbf{R}^m$ and $g : X \to \mathbf{R}^n$ are both continuous if and only if $f \oplus g : X \to \mathbf{R}^{m+n}$ is continuous. $\square$

Exercise 2.2.7. *Let $k \geq 1$, let I be a finite subset of $\mathbf{N}^k$, and let $c : I \to \mathbf{R}$ be a function. Form the function $P : \mathbf{R}^k \to \mathbf{R}$ defined by*

$$P(x_1, \cdots, x_k) := \sum_{(i_1, \cdots, i_k) \in I} c(i_1, \cdots, i_k) x^{i_1} \cdots x^{i_k}.$$

(Such a function is known as a polynomial *of k variables; a typical example of such a polynomial is $P(x_1, x_2, x_3) = 3x_1^3 x_2 x_3^2 - x_2 x_3^2 + x_1 + 5$.) Show that P is continuous. (Hint: use induction on k, Exercise 2.2.6, and either Exercise 2.2.5 or Lemma 2.2.2.)*

Proof. Since c does not make sense, we only consider

$$x^{i_1} \cdots x^{i_k}.$$

Use induction on k. Base case $k = 1$ is obviously true by Lemma 2.2.2. Inductively suppose that assertion is hold for k, i.e., $x^{i_1} \cdots x^{i_k}$ is continuous. By Lemma 2.2.2,

$$x^{i_1} \cdots x^{i_k} x^{i_{k+1}}$$

is also continuous. This close the induction. Use Lemma 2.2.2 repeatedly, we can conclude that

$$P(x_1, \cdots, x_k) := \sum_{(i_1, \cdots, i_k) \in I} c(i_1, \cdots, i_k) x^{i_1} \cdots x^{i_k}.$$

is continuous. In general, if $f_i : X \to \mathbf{R}$ are continuous for all $i = 1, \cdots, k$. Then $f_1 \oplus \cdots \oplus f_k : X \to \mathbf{R}^k$ also continuous. Then $P(f_1, \cdots, f_k) = (f_1 \oplus \cdots \oplus f_k) \circ P$ is continuous. $\square$

Exercise 2.2.8. *Let (X, d_X) and (Y, d_Y) be metric spaces. Define the metric $d_{X \times Y} : (X \times Y) \times (X \times Y) \to [0, \infty)$ by the formula*

$$d_{X \times Y}((x, y), (x', y')) := d_X(x, x') + d_Y(y, y').$$

Show that $(X \times Y, d_{X \times Y})$ is a metric space, and deduce an analogue of Proposition 1.1.18 and Lemma 2.2.1.

Proof.

(a) For any $(x, y) \in X \times Y$, we have

$$d_{X \times Y}((x, y), (x, y)) = d_X(x, x) + d_Y(y, y) = 0.$$

(b) For any distinct $(x, y), (x', y') \in X \times Y$, we have

$$d_{X \times Y}((x, y), (x', y')) = d_X(x, x') + d_Y(y, y') \geq 0.$$

(c) For any $(x, y), (x', y') \in X \times Y$, we have

$$\begin{aligned}
d_{X \times Y}((x, y), (x', y')) &= d_X(x, x') + d_Y(y, y') \\
&= d_X(x', x) + d_Y(y', y) \\
&= d_{X \times Y}((x', y'), (x, y)).
\end{aligned}$$

(d) For any $(x, y), (x', y'), (x'', y'') \in X \times Y$, we have

$$\begin{aligned}
d_{X \times Y}((x, y), (x'', y'')) &= d_X(x, x'') + d_Y(y, y'') \\
&\leq d_X(x, x') + d_X(x', x'') + d_Y(y, y') + d_Y(y', y'') \\
&= d_{X \times Y}((x, y), (x', y')) + d_{X \times Y}((x', y'), (x'', y'')).
\end{aligned}$$

Thus $(X \times Y, d_{X \times Y})$ is a metric space. $\qquad\square$

Proposition 1. Let $(X \times Y, d_{X \times Y})$ be a metric space, and let $(x^{(n)})_{n=m}^{\infty}$ and $(y^{(n)})_{n=m}^{\infty}$ be two sequences of points in X and Y respectively. Let $(x^{(n)}, y^{(n)})_{n=m}^{\infty}$ be a sequence of points in $X \times Y$. Then the following two statements are equivalent:

(a) $(x^{(n)}, y^{(n)})_{n=m}^{\infty}$ converges to (x, y) with respect to the metric $d_{X \times Y}$.

(b) $(x^{(n)})_{n=m}^{\infty}$ converges to x with respect to the metric d_X and $(y^{(n)})_{n=m}^{\infty}$ converges to y with respect to the metric d_Y.

Proof of Proposition 1. (a) $\Rightarrow$ (b). Suppose that $(x^{(n)}, y^{(n)})_{n=m}^{\infty}$ converges to (x, y) with respect to the metric $d_{X \times Y}$. This means that for every $\varepsilon > 0$ there exists an $N \geq m$ such that

$$d_{X \times Y}((x^{(n)}, y^{(n)}), (x, y)) = d_X(x^{(n)}, x) + d_Y(y^{(n)}, y) \leq \varepsilon$$

for all $n \geq N$. Since $d_X(x^{(n)}, x) \geq 0$ and $d_Y(y^{(n)}, y) \geq 0$, for every $\varepsilon > 0$ there exists an $N \geq m$ such that $d_X(x^{(n)}, x) \leq \varepsilon$ and $d_Y(y^{(n)}, y) \leq \varepsilon$ for all $n \geq N$. Thus $(x^{(n)})_{n=m}^{\infty}$ converges to x and $(y^{(n)})_{n=m}^{\infty}$ converges to y.

(b) $\Rightarrow$ (a). Suppose that $(x^{(n)})_{n=m}^{\infty}$ converges to x and $(y^{(n)})_{n=m}^{\infty}$ converges to y. This means that for every $\varepsilon > 0$ there exists an $N' \geq m$ and $N'' \geq m$ such that

$$d_X(x^{(n)}, x) \leq \varepsilon, \quad \text{and} \quad d_Y(y^{(n)}, y) \leq \varepsilon$$

for all $n \geq N'$ and $n \geq N''$. Let $N = \max(N', N'')$. then for all $\zeta = 2\varepsilon > 0$ there exists an $N \geq m$ such that

$$d_{X \times Y}((x^{(n)}, y^{(n)}), (x, y)) = d_X(x^{(n)}, x) + d_Y(y^{(n)}, y) \leq \zeta$$

for all $n \geq N$. Thus $(x^{(n)}, y^{(n)})_{n=m}^{\infty}$ converges to (x, y). $\qquad\square$

> **Lemma 1.** Let $f : X \to U$ and $g : X \to V$ be functions, and let $f \oplus g : X \to U \times V$ be their direct sum. We give $(U \times V, d_{U \times V})$ be a metric space.
>
> (a) If $x_0 \in X$, then f and g are both continuous at x_0 if and only if $f \oplus g$ is continuous at x_0.
>
> (b) f and g are both continuous if and only if $f \oplus g$ is continuous.

Proof of Lemma 1. (a) Since f and g are continuous at x_0, by Theorem 2.1.4, for all $(x^{(n)})_{n=1}^{\infty}$ in X converges to x_0, $(f(x^{(n)}))_{n=1}^{\infty}$ and $(g(x^{(n)}))_{n=1}^{\infty}$ converges to $f(x_0)$ and $g(x_0)$ respectively. Let $(f(x^{(n)}), g(x^{(n)}))_{n=1}^{\infty}$ be a sequence in $U \times V$. By Proposition 1, $(f(x^{(n)}), g(x^{(n)}))_{n=1}^{\infty}$ converges to $(f(x_0), g(x_0))$. Thus $f \oplus g$ is continuous at x_0. While if $f \oplus g$ is continuous at x_0, by Theorem 1.1.4, for every $(x^{(n)})_{n=1}^{\infty}$ in X converges to x_0, $(f(x^{(n)}), g(x^{(n)}))_{n=1}^{\infty}$ converges to $(f(x_0), g(x_0))$. By Proposition 1, $(f(x^{(n)}))_{n=1}^{\infty}$ and $(g(x^{(n)}))_{n=1}^{\infty}$ converges to $f(x_0)$ and $g(x_0)$ respectively. Thus f and g are continuous.

(b) is immediately follows from (a). $\qquad\square$

Exercise 2.2.9. *Let $f : \mathbf{R}^2 \to \mathbf{R}$ be a function from $\mathbf{R}^2$ to $\mathbf{R}$. Let (x_0, y_0) be a point in $\mathbf{R}^2$. If f is continuous at (x_0, y_0), show that*

$$\lim_{x \to x_0} \limsup_{y \to y_0} f(x, y) = \lim_{y \to y_0} \limsup_{x \to x_0} f(x, y) = f(x_0, y_0)$$

and

$$\lim_{x \to x_0} \liminf_{y \to y_0} f(x, y) = \lim_{y \to y_0} \liminf_{x \to x_0} f(x, y) = f(x_0, y_0).$$

(Recall that $\limsup_{x \to x_0} f(x) := \inf_{r>0} \sup_{|x-x_0|<r} f(x)$ and $\liminf_{x \to x_0} f(x) := \sup_{r>0} \inf_{|x-x_0|<r} f(x)$).) In particular, we have

$$\lim_{x \to x_0} \lim_{y \to y_0} f(x, y) = \lim_{y \to y_0} \lim_{x \to x_0} f(x, y)$$

whenever the limits on both sides exist. (Note that the limits do not necessarily exist in general; consider for instance the function $f : \mathbf{R}^2 \to \mathbf{R}$ such that $f(x, y) = y \sin \frac{1}{x}$ when $xy \neq 0$ and $f(x, y) = 0$ otherwise.) Discuss the comparison between this result and Example 1.2.7.

Proof. Since f is continuous at (x_0, y_0), for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $|f(x, y) - f(x_0, y_0)| < \varepsilon$ whenever $|(x, y) - (x_0, y_0)| < \delta$. Given $y' \in \mathbf{R}$, by Exercise 2.2.4, $f(x, y') = f_1(x)$ is continuous. Then for every $\varepsilon_1 > 0$ there exists a $\delta_1 > 0$ such that $|f_1(x) - f_1(x_0)| < \varepsilon_1$ whenever $|(x, y') - (x_0, y)| < \delta_1$. Thus $\limsup_{x \to x_0} f(x, y') =$

<div style="border:1px solid;">

2.3

Continuity and compactness

</div>

Exercise 2.3.1. *Prove Theorem 2.3.1.*

> **Theorem 2.3.1** (Continuous maps preserve compactness). Let $f : X \to Y$ be a continuous map from one metric space (X, d_X) to another (Y, d_Y). Let $K \subseteq X$ be any compact subset of X. Then the image $f(K) := \{f(x) : x \in K\}$ of K is also compact.

Proof. Suppose that $\bigcup_{\alpha \in I} V_\alpha$ be an open cover of $f(K)$. Since f is continuous, by Theorem 2.1.5, $f^{-1}(\bigcup_{\alpha \in I} V_\alpha)$ is an open cover of K. Since K is compact, there is a finite open cover $f^{-1}(\bigcup_{\alpha \in F} V_\alpha)$ of K. Thus $\bigcup_{\alpha \in F} V_\alpha$ is

a finite open cover of $f(K)$. By the converse of Theorem 1.5.8,[2] $f(K)$ is compact. $\qquad\square$

Exercise 2.3.2. *Prove Proposition 2.3.2. (Hint: modify the proof of Proposition 9.6.7.)*

> **Proposition 2.3.2** (Maximum principle). Let (X, d) be a compact metric space, and let $f : X \to \mathbf{R}$ be a continuous function. Then f is bounded[3]. Furthermore, if X is non-empty, f attains its maximum at some point $x_{max} \in X$, and also attains its minimum at some point $x_{min} \in X$.

Proof. Since X is compact, for all $\varepsilon > 0$ there exists finite balls such that

$$X = \bigcup_{i=1}^{n} B(x^{(i)}, \varepsilon).$$

Then

$$f(X) = f\left(\bigcup_{i=1}^{n} B(x^{(i)}, \varepsilon)\right).$$

Since f is continuous, we have

$$f\left(\bigcup_{i=1}^{n} B(x^{(i)}, \varepsilon)\right) \subseteq \bigcup_{i=1}^{n} B(f(x^{(i)}), \varepsilon).$$

By Heine-Borel theorem, $f(X)$ is compact. Thus $f(X)$ is bounded.

Let $S := \sup_{x \in X} f(x)$. By Exercise 1.5.10, X is complete, then there exists a subsequence $(x^{n_j})_{j=m}^{\infty}$ which converges to x_{max}. Since f is continuous, when $(x^{n_j})_{j=m}^{\infty}$ converges to x_{max}, we have $(f(x^{n_j}))_{j=m}^{\infty}$ converges to S. Thus f attains its maximum. This is similar to show that f attains its minimum. $\qquad\square$

Exercise 2.3.3. *Show that every uniformly continuous function is continuous, but give an example that shows that not every continuous function is uniformly continuous.*

---

[2]See *Exercise 1.5.11.*

[3]It should be "compact" other than "bounded".

Proof. Suppose f is uniformly continuous, by Definition 2.3.4, we given x_0 for all $x_0 \in X$, then obtained that f is continuous.

For the converse, the function $f : \mathbf{R}^2 \to \mathbf{R}^2$ defined by $f(x, y) := (\frac{1}{x}, \frac{1}{y})$ is not uniformly continuous. $\square$

Exercise 2.3.4. *Let (X, d_X), (Y, d_Y), (Z, d_Z) be metric spaces, and let $f : X \to Y$ and $g : Y \to Z$ be two uniformly continuous functions. Show that $g \circ f : X \to Z$ is also uniformly continuous.*

Proof. Suppose that f and g are uniformly continuous. By Definition 2.3.4, for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $d_Y(f(x), f(x')) < \varepsilon$ whenever $x, x' \in X$ are such that $d_X(x, x') < \delta$; for every $\zeta > 0$ there exists a $\varepsilon > 0$ such that $d_Z(g(y), g(y')) < \zeta$ whenever $y, y' \in Y$ are such that $d_Y(y, y')$. Thus for every $\zeta > 0$ there exists a $\delta > 0$ such that $d_Z(g(f(x)), g(f(x'))) < \zeta$ whenever $y = f(x), y' = f(x') \in Y$ are such that $d_X(x, x') < \delta$. Thus $g \circ f$ is uniformly continuous. $\square$

Exercise 2.3.5. *Let (X, d_X) be a metric space, and let $f : X \to \mathbf{R}$ and $g : X \to \mathbf{R}$ be uniformly continuous functions. Show that the direct sum $f \oplus g : X \to \mathbf{R}^2$ defined by $f \oplus g(x) := (f(x), g(x))$ is uniformly continuous.*

Proof. Suppose that f and g are uniformly continuous. By Definition 2.3.4, for all $\varepsilon > 0$ there exists a $\delta > 0$ such that $d_X(f(x), f(x')) < \varepsilon$ and $d_X(g(x), g(x')) < \varepsilon$ whenever $x, x' \in X$ are such that $d(x, x') < \delta$. Then we have $d(f \oplus g(x), f \oplus g(x')) < 2\varepsilon$ whenever $x, x' \in X$ are such that $d(x, x') < \delta$. Thus $f \oplus g$ is uniformly continuous. $\square$

Exercise 2.3.6. *Show that the addition function $(x, y) \mapsto x + y$ and the subtraction function $(x, y) \mapsto x - y$ are uniformly continuous from $\mathbf{R}^2$ to $\mathbf{R}$, but the multiplication function $(x, y) \mapsto xy$ is not. Conclude that if $f : X \to \mathbf{R}$ and $g : X \to \mathbf{R}$ are uniformly continuous functions on a metric space (X, d), then $f + g : X \to \mathbf{R}$ and $f - g : X \to \mathbf{R}$ are also uniformly continuous. Give an example to show that $fg : X \to \mathbf{R}$ need not be uniformly continuous. What is the situation for $\max(f, g)$, $\min(f, g)$,*

f/g, and cf for a real number c?

Proof. By Definition 2.3.4, we need to show that for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d_{\mathbf{R}}(x + y, x' + y') < \varepsilon$ whenever $(x, y), (x', y') \in \mathbf{R}^2$ are such that $d_{\mathbf{R}^2}((x, y), (x', y')) < \delta$.

2.4 Continuity and connectedness

2.5 Topological spaces (Optional)

Exercise 2.5.1. *Let X be an arbitrary set, and let $\mathcal{F} := \{\emptyset, X\}$. Show that $(X, \mathcal{F})$ is a topology (called the* trivial topology *on X). If X contains more than one element, show that the trivial topology cannot be obtained from by placing a metric d on X. Show that this topological space is both compact and connected.*

Proof. We can see that empty set $\emptyset$ and whole set X lie in $\mathcal{F}$, and $X \cap \emptyset = \emptyset \in \mathcal{F}$, $X \cup \emptyset = X \in \mathcal{F}$. By Definition 2.5.1, $(X, \mathcal{F})$ is a topology.

Suppose that $X := \{a, b\}$. Let (X, d) to be a metric space. Since $\{a\}$ and $\{b\}$ are closed in X, $X = \{a\} \cup \{b\}$ is also closed in X. Thus $\mathcal{F}$ is not a topological space.

Let $(\mathcal{O}_\alpha)_{\alpha \in I}$ be any open cover of X, i.e., $X \subseteq \bigcup_{\alpha \in I} \mathcal{O}_\alpha$. We can see that there is a finite open cover $\mathcal{F}$ of X such that $X \subseteq \emptyset \cup X$. Thus X is compact.

In the topological space $(X, \mathcal{F})$, X is not disconnected for that we only have $X = \emptyset \cup X$. Thus X is connected. $\square$

Exercise 2.5.2. *Let (X, d) be a metric space (and hence a topological space). Show that the two notions of convergence of sequences in Definition 1.1.14 and Definition 2.5.4 coincide.*

Proof. Let $(x^{(n)})_{n=m}^{\infty}$ be a sequence of points in X which converges to x in X. Let (X, d) be a metric space. Suppose in the sense of Definition 1.1.14, for every $\varepsilon > 0$ there exists an $N \geq m$ such that $d(x^{(n)}, x) < \varepsilon$ for all $n \geq N$, so that $x^{(n)} \in B(x, \varepsilon)$ for all $n \geq N$. Let V be any neighbourhood of x in $\mathcal{F}$. Let V be any neighbourhood of x in $\mathcal{F}$. Then for every V, there is a $B(x, \varepsilon) \subseteq V$. Thus for every neighbourhood V of x, there is an $N \geq m$ such that $x^{(n)} \in V$ for all $n \geq N$.

For the converse, suppose in the sense of Definition 2.5.4, for every neighbourhood V of x, there is an $N \geq m$ such that $x^{(n)} \in V$ for all $n \geq N$. Since $B(x, \varepsilon)$ is a neighbourhood of x for all ε. Thus for every $\varepsilon > 0$, $B(x, r)$ is a neighbourhood of x, there exists an $N \geq m$ such that $x^{(n)} \in B(x, \varepsilon)$, so that $d(x^{(n)}, x) < \varepsilon$ for all $n \geq N$. $\square$

Exercise 2.5.3. *Let (X, d) be a metric space (and hence a topological space). Show that the two notions of interior, exterior, and boundary in Definition 1.2.5 and Definition 2.5.5 coincide.*

Proof. Let E be a subset of x, let x_0 be a point in E. Let (X, d) be a metric space and suppose in the sense of Definition 1.2.5. If x_0 is an interior point of E, there is a radius $r > 0$ such that $B(x_0, r) \subseteq E$. Since $B(x_0, r)$ is a neighbourhood of x_0, x_0 is also an interior point of E in $(X, \mathcal{F})$. A similar argument shows that if x_0 is an exterior or a boundary point of E in (X, d) implies that it also is an exterior or a boundary point of E in $(X, \mathcal{F})$.

Conversely, let $(X, \mathcal{F})$ be a topological space and suppose in the sense of Definition 2.5.5. If x_0 is an interior point of E, there exists a neighbourhood of x_0 such that $V \subseteq E$. Then there is a ball $B(x_0, r) \in V$ such that $B(x_0, r) \subseteq E$. Thus x_0 is an interior point of E in (X, d). A similar argument shows that if x_0 is an exterior or a boundary point of E in $(X, \mathcal{F})$ implies that it also is an exterior or a boundary point of E in (X, d). $\square$

Exercise 2.5.4. *A topological space $(X, \mathcal{F})$ is said to be* Hausdorff *if given any two distinct points $x, y \in X$, there exists a neighbourhood V of x and a neighbourhood W of y such that $V \cap W = \emptyset$. Show that any topological space*

coming from a metric space is Hausdorff, and show that the trivial topology is not Hausdorff. Show that the analogue of Proposition 1.1.20 holds for Hausdorff topological spaces, but give an example of a non-Hausdorff topological space in which Proposition 1.1.20 fails. (In practice, most topological spaces one works with are Hausdorff; non-Hausdorff topological spaces tend to be so pathological that it is not very profitable to work with them.)

Proof. Let (X, d) be a metric space. Let $(x^{(n)})_{n=m}^{\infty}$ and $(y^{(n)})_{n=m}^{\infty}$ take points in X and converge to x and y respectively, where $x \neq y$. By Definition 2.5.4, for every neighbourhood V of x, there exists an $N \geq m$ such that $x^{(n)} \in V$ for all $n \geq N$. By Proposition 1.1.20, for all $N \geq m$, there exists a neighbourhood V of x such that $y^{(n)} \notin V$ for all $n \geq N$. For such a neighbourhood V of x, we can find a neighbourhood W of y such that $y^{(n)} \in W$ for all $n \geq M$. Takes $\max(N, M)$, we have $W \cap V = \emptyset$.

Suppose trivial topology $(X, \mathcal{F})$. The only neighbourhood of x and y is X. Thus X is not Hausdorff. $\qquad\square$

Now we show that the following Proposition holds for Hausdorff topological spaces:

> **Proposition 1.** Let $(X, \mathcal{F})$ be a topological space with Hausdorff, and let $(x^{(n)})_{n=m}^{\infty}$ be a sequence in X. Suppose that there are two points $x, x' \in X$ such that $(x^{(n)})_{n=m}^{\infty}$ converges to x, and $(x^{(n)})_{n=m}^{\infty}$ also converges to x'. Then we have $x = x'$.

Proof of Proposition 1. Suppose for sake of contradiction that $x \neq x'$, by Hausdorff, there is a neighbourhood V of x and a neighbourhood W of x' such that $V \cap W = \emptyset$. On the one hand, $(x^{(n)})_{n=m}^{\infty}$ converges to x, then for V above, there exists an $N \geq m$ such that $x^{(n)} \in V$ for all $n \geq N$. Other hand, $(x^{(n)})_{n=m}^{\infty}$ converges to x', then for neighbourhood W above, there exists an $M \geq m$ such that $x^{(n)} \in W$ for all $n \geq M$. Takes $max(N, M)$, we have $x^{(n)} \in V$ and $x^{(n)} \in W$ at the time, a contradiction.

If X is not Hausdorff, for example, $\mathcal{F} = \{\emptyset, X\}$. Let $(X, \mathcal{F})$ be a trivial topology which is not Hausdorff. Let $(x^{(n)})_{n=m}^{\infty}$ be a sequence in X. Then for any two distinct points $x, x' \in X$, $(x^{(n)})_{n=m}^{\infty}$ converges to x and x' at the time, but $x \neq x'$. $\qquad\square$

Exercise 2.5.5. *Given any totally ordered set X with order relation $\leq$, declare a set $V \subset X$ to be* open *if for every $x \in V$ there exists a set I which is an interval $\{y \in X : a < y < b\}$ for some $a, b \in X$, a ray $\{y \in X : a < y\}$ for some $a \in X$, the ray $\{y \in X : y < b\}$ for some $b \in X$, or the whole space X, which contains x and is contained in V. Let $\mathcal{F}$ be the set of all open subsets of X. Show that $(X, \mathcal{F})$ is a topology (this is the* order topology *on the totally ordered set $(X, \leq)$) which is Hausdorff in the sense of Exercise 2.5.4. Show that on the real line $\mathbf{R}$ (with the standard ordering $\leq$), the order topology matches the standard topology (i.e., the topology arising from the standard metric). If instead one applies this to the extended real line $\mathbf{R}^*$, show that $\mathbf{R}$ is an open set with boundary $\{-\infty, +\infty\}$. If $(x_n)_{n=1}^{\infty}$ is a sequence of numbers in $\mathbf{R}$ (and hence in $\mathbf{R}^*$), show that x_n converges to $+\infty$ if and only if $\liminf_{n\to\infty} x_n = +\infty$, and x_n converges to $-\infty$ if and only if $\limsup_{n\to\infty} x_n = -\infty$.*

Proof. We can see that $\emptyset$ satisfies definition of openness vacuously, hence we have $\emptyset \in \mathcal{F}$. Since for every $x \in X$, there is some I containing x and $I \subseteq X$, we have $X \in \mathcal{F}$.

Let $(V_\alpha)_{\alpha \in A}$ be the elements of $\mathcal{F}$, and we need to show that $V = \bigcup_{\alpha \in A} V_\alpha \in \mathcal{F}$. Given $x \in V$, there is an α such that $x \in V_\alpha$. By definition of openness, there is an I such that $x \in I \subseteq V_\alpha$. Then $x \in I$ and $I \subseteq V$, so that V is open and $V \in \mathcal{F}$, by definition.

To show that $V_1 \cap \cdots \cap V_n$ is contained in $\mathcal{F}$, we first show that $V_1 \cap V_2 \in \mathcal{F}$ for $V_1, V_2 \in \mathcal{F}$. Given $x \in V_1 \cap V_2$ (otherwise, $V_1 \cap V_2 = \emptyset$ and belongs to $\mathcal{F}$), there is I_1, I_2 such that $x \in I_1 \subseteq V_1$ and $x \in I_2 \subseteq V_2$. We consider following different cases:

(1) If $I_1 = \{y \in X : a_1 < y\}, I_2 = \{y \in X : a_2 < y\}$, then there is an $I_3 = \{y \in X : a < y\}$ for $a = \max(a_1, a_2)$ such that $I_3 \subseteq V_1 \cap V_2$ and $x \in I_3$.

(2) If $I_1 = \{y \in X : y < b_1\}, I_2 = \{y \in X : y < b_2\}$, then there is an $I_3 = \{y \in X : y < b\}$ for $b = \min(b_1, b_2)$ such that $I_3 \subseteq V_1 \cap V_2$ and $x \in I_3$.

(3) If $I_1 = \{y \in X : a_1 < y < b_1\}, I_2 = \{y \in X : a_2 < y < b_2\}$, then there is an $I_3 = \{y \in X : a < y < b\}$ for $a = \max(a_1, a_2), b = \min(b_1, b_2)$ such that $I_3 \subseteq V_1 \cap V_2$ and $x \in I_3$.

(4) Remaining cases is similar.

We can conclude that there is an $I_3 \subseteq I_1 \cap I_2$ such that $I_3 \subseteq V_1 \cap V_2$ and $x \in I_3$. Thus $V_1 \cap V_2 \in \mathcal{F}$.

Now we use induction on n. The base case $n = 1$ is trivial. Suppose inductively that $(V_1 \cap \cdots \cap V_{n-1}) \in \mathcal{F}$, then there is a subset I_1 containing x. Let V_n belongs to $\mathcal{F}$, there is an $I_2 \subseteq V_n$. By the conclusion above, if $(V_1 \cap \cdots \cap V_{n-1}) \cap V_n$ is not empty, there is an $I_3 \subseteq I_1 \cap I_2$ such that $I_3 \subseteq V_1 \cap \cdots \cap V_n$ and $x \in I_3$. Thus $(V_1 \cap \cdots \cap V_n) \in \mathcal{F}$.

Thus $(X, \mathcal{F})$ is a topology.

---

[3]Following materials from J. R. Munkres is useful:

> **Definition.** If X is a set, a *basis* for a topology on X is a collection $\mathcal{B}$ of subsets of X, called *basis elements*, such that
>
> (a) For each $x \in X$, there is at least one basis element B containing x.
>
> (b) If x belongs to the intersection of two basis elements B_1 and B_2, then there is a basis element B_3 containing x such that $B_3 \subseteq B_1 \cap B_2$.

If $\mathcal{B}$ satisfies these two conditions, then we define the *topology $\mathcal{F}$ generated by $\mathcal{B}$* as follows: *A subset V of X is said to be open in X (that is, to be an element of $\mathcal{F}$) if for each $x \in V$, there is a basis element $B \in \mathcal{B}$ such that $x \in B$ and $B \subseteq V$.* Note that each basis element is itself an element of $\mathcal{F}$.

> **Proposition.** The collection $\mathcal{F}$ generated by the basis $\mathcal{B}$ is a topology on X.

Proof. If V is the empty set, it satisfies the defining condition of openness vacuously. Likewise, X is in $\mathcal{F}$, since for each $x \in X$ there is some basis element B containing x and contained in X. Now let us take an indexed family $\{V_\alpha\}_{\alpha \in I}$, of elements of $\mathcal{F}$ and show that

$$V = \bigcup_{\alpha \in J} V_\alpha$$

belongs to $\mathcal{F}$. Given $x \in V$, there is an index α such that $x \in V_\alpha$. Since V_α is open, there is a basis element B such that $x \in B \subseteq V_\alpha$. Then $x \in B$ and $B \subseteq V$, so that V is open, by definition.

We show that order topology $(X, \mathcal{F})$ is Hausdorff. Let x and y be any two distinct points from X. We must only have $x < y$ or $y < x$. Suppose $x < y$. Consider the set

$$I := \{z \in X : x < a < y\}.$$

If I is empty, then x is contained in neighbourhood $\{a \in X : a < y\}$ and y is contained in neighbourhood $\{a \in X : x < a\}$, where

$$\{a \in X : a < y\} \cap \{a \in X : x < a\} = \emptyset.$$

While if I is not empty, then x is contained in neighbourhood $\{b \in X : b < a\}$ and y is contained in neighbourhood $\{b \in X : a < b\}$, where

$$\{b \in X : b < a\} \cap \{b \in X : a < b\} = \emptyset.$$

Thus X is Hausdorff.

Now let us take two elements V_1 and V_2 of $\mathcal{F}$ and show that $V_1 \cap V_2$ belongs to $\mathcal{F}$. Given $x \in V_1 \cap V_2$, choose a basis element B_1 containing x such that $B_1 \subseteq V_1$; choose also a basis element B_2 containing x such that $B_2 \subseteq V_2$. The second condition for a basis enables us to choose a basis element B_3 containing x such that $B_3 \subseteq B_1 \cap B_2$. Then $x \in B_3$ and $B_3 \subseteq V_1 \cap V_2$, so $V_1 \cap V_2$ belongs to $\mathcal{F}$, by definition.

Finally, we show by induction that any finite intersection $V_1 \cap \cdots \cap V_n$ of elements of $\mathcal{F}$ is in $\mathcal{F}$. This fact is trivial for $n = 1$; we suppose it true for $n - 1$ and prove it for n. Now

$$(V_1 \cap \cdots \cap V_n) = (V_1 \cap \cdots \cap V_{n-1}) \cap V_n.$$

By hypothesis, $V_1 \cap \cdots \cap V_{n-1}$ belongs to $\mathcal{F}$; by the result just proved, the intersection of $V_1 \cap \cdots \cap V_{n-1}$ and V_n also belongs to F.

Thus we have checked that collection of open sets generated by a basis $\mathcal{B}$ is, in fact, a topology. $\square$

Uniform convergence

Limiting values of functions

Exercise 3.1.1. *Let (X, d_X) and (Y, d_Y) be metric spaces, let E be a subset of X, let $f : E \to Y$ be a function, and let x_0 be an element of E. Show that the limit $\lim_{x \to x_0; x \in E} f(x)$ exists if and only if the limit $\lim_{x \to x_0; x \in E \setminus \{x_0\}} f(x)$ exists and is equal to $f(x_0)$. Also, show that if the limit $\lim_{x \to x_0; x \in E} f(x)$ exists at all, then it must equal $f(x_0)$.*

Proof. Suppose that the limit $\lim_{x \to x_0; x \in E} f(x)$ exists and equals to $f(x_0)$. By Definition 3.1.1, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d_Y(f(x), f(x_0)) < \varepsilon$ for all $x \in E$ such that $d_X(x, x_0) < \delta$. This implies that $d_Y(f(x), f(x_0)) < \varepsilon$ for all $x \in E \setminus \{x_0\}$ such that $0 < d_X(x, x_0) < \delta$. Thus by definition, $\lim_{x \to x_0; x \in E \setminus \{x_0\}} f(x) = f(x_0)$.

For the converse, suppose that $\lim_{x \to x_0; x \in E \setminus \{x_0\}} f(x) = f(x_0)$. By Definition 3.1.1, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d_Y(f(x), f(x_0)) < \varepsilon$ for all $x \in E \setminus \{x_0\}$ such that $0 < d_X(x, x_0) < \delta$. When $x = x_0$, we have $d(x, x_0) = 0$. Thus we have $d_Y(f(x), f(x_0)) < \varepsilon$ for all $x \in E$ such that

$d_X(x, x_0) < \delta$.

Finally, we suppose for sake of contradiction that $\lim_{x \to x_0; x \in E} f(x)$ exists and it not equal $f(x_0)$. Let $\lim_{x \to x_0; x \in E} f(x) = L$. By Definition 3.1.1, for every $\varepsilon > 0$ there is a $\delta > 0$ such that $d_Y(f(x), L) < \varepsilon$ for all $x \in E$ such that $d_X(x, x_0) < \delta$. Because there exists an $x = x_0$ such that $d_Y(f(x_0), L) > \zeta$ where $\zeta > 0$ such that $d_X(x, x_0) < \delta$. Let $\varepsilon = \zeta/2$, then $d_Y(f(x), L) > \varepsilon$ for $x_0 \in E$ such that $d_X(x_0, x_0) < \delta$, a contradiction. Thus $\lim_{x \to x_0; x \in E} f(x)$ must equals to $f(x_0)$. □

Exercise 3.1.2. *Prove Proposition 3.1.5. (Hint: review your proof of Theorem 2.1.4.)*

> **Proposition 3.1.5.** Let (X, d_X) and (Y, d_Y) be metric spaces, let E be a subset of X, and let $f : X \to Y$ be a function. Let $x_0 \in X$ be an adherent point of E and $L \in Y$. Then the following four statements are logically equivalent:
>
> (a) $\lim_{x \to x_0; x \in E} f(x) = L$.
>
> (b) For every sequence $(x^{(n)})_{n=1}^\infty$ in E which converges to x_0 with respect to the metric d_X, the sequence $(f(x^{(n)}))_{n=1}^\infty$ converges to L with respect to the metric d_Y.
>
> (c) For every open set $V \subset Y$ which contains L, there exists an open set $U \subset X$ containing x_0 such that $f(U \cap E) \subseteq V$.
>
> (d) If one defines the function $g : E \cup \{x_0\} \to Y$ by defining $g(x_0) := L$, and $g(x) := f(x)$ for $x \in E \setminus \{x_0\}$, then g is continuous at x_0. Furthermore, if $x_0 \in E$, then $f(x_0) = L$.

Proof. (a) $\Rightarrow$ (b). By Definition 3.1.1, for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $d_Y(f(x), L) < \varepsilon$ for all $x \in E$ such that $d_X(x, x_0) < \delta$. Since every sequence $(x^{(n)})_{n=1}^\infty$ in E which converges to x_0 with respect to the metric d_X, by Definition 1.1.14, for every $\delta > 0$, there exists an $N \geq 1$ such that $d_X(x, x_0) \leq \delta$ for all $n \geq N$. Thus for every $\varepsilon > 0$ there exists an $N \geq 1$

such that $d_Y(f(x^{(n)}), L) \leq \varepsilon$ for all $n \geq N$. By Definition 1.1.14 again, the sequence $(f(x^{(n)}))_{n=1}^\infty$ converges to L with respect to the metric d_Y.

(b) $\Rightarrow$ (c). Since V is open in Y which contains L, by Proposition 1.2.15(a), there is a $r > 0$ such that $B(L, r) \subseteq V$. Let $\varepsilon = r/2$, for every $\delta > 0$ there exists an $N \geq 1$ such that $d_Y(f(x^{(n)}), L) < \varepsilon$ for all $n \geq N$ whenever $d_X(x, x_0) \leq \delta$, so that $d_X(x, x_0) < \delta$. Thus there exists $U = B(x_0, \delta) \subset X$ containing x_0 such that $f(U \cap E) \subseteq V$.

(c) $\Rightarrow$ (b). For every $\varepsilon > 0$ and $\delta > 0$, $B(L, \varepsilon)$ and $B(x_0, \delta)$ are open. Then for every $\varepsilon > 0$ such that $B(L, \varepsilon) \subset Y$, there exists a $\delta > 0$ such that $B(x_0, \delta) \subset X$ such that $f(B(x_0, \delta) \cap E) \subseteq B(L, \varepsilon)$. This means that for every $\varepsilon > 0$ such that $d_Y(f(x), L) < \varepsilon$ for all $x \in E$, there is a $\delta > 0$ such that $d_X(x, x_0) < \delta$. Thus $\lim_{x \to x_0; x \in E} f(x) = L$.

(a) $\Leftrightarrow$ (d). Suppose that $f(x)$ converges to L as x converges to x_0. Then for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $d_Y(f(x), L) = d_Y(g(x), g(x_0)) < \varepsilon$ for all $x \in E \setminus \{x_0\}$ such that $d_X(x, x_0) < \delta$. Thus g is continuous at x_0. Furthermore, if $x_0 \in E$, we have $g(x_0) = f(x_0) = L$. Conversely, suppose that g is continuous at x_0. By Definition 2.1.1, for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $d_Y(g(x), g(x_0)) = d_Y(f(x), L)$ whenever $d_X(x, x_0) < \delta$. We can see that this is hold for, if $x_0 \notin E$, we have $\lim_{x \to x_0; x \in E \setminus \{x_0\}} f(x) = L$, and if $x_0 \in E$, then $\lim_{x \to x_0; x \in E} f(x) = L$. Thus $f(x_0)$ is converges to L as x converges to x_0. $\qquad \square$

<div style="border:1px solid;">

3.2

Pointwise and uniform convergence

</div>

Exercise 3.2.1. *The purpose of this exercise is to demonstrate a concrete relationship between continuity and pointwise convergence, and between uniform continuity and uniform convergence. Let $f : \mathbf{R} \to \mathbf{R}$ be a function. For any $a \in \mathbf{R}$, let $f_a : \mathbf{R} \to \mathbf{R}$ be the shifted function $f_a(x) := f(x - a)$.*

(a) *Show that f is continuous if and only if, whenever $(a_n)_{n=0}^\infty$ is a sequence of real numbers which converges to zero, the shifted functions*

f_{a_n} converge pointwise to f.

(b) *Show that f is uniformly continuous if and only if, whenever $(a_n)_{n=0}^{\infty}$ is a sequence of real numbers which converges to zero, the shifted functions f_{a_n} converge uniformly to f.*

Proof. (a) Suppose that f is continuous, by Definition 2.1.1, for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $d(f(x), f(x')) < \varepsilon$ whenever $d(x, x') < \delta$ for all $x, x' \in \mathbf{R}$. When $(a_n)_{n=0}^{\infty}$ converges to 0, for every $\delta > 0$ there is an $N > 0$ such that $d(a_n, 0) < \delta$ for all $n \geq N$. Therefore, for every $\varepsilon > 0$ and $x \in \mathbf{R}$, there exists an $N > 0$ such that $d(f_{a_n}(x) - f(x)) = d(f(x - a_n), f(x)) < \varepsilon$ for every $n > N$, whenever $d(x - a_n, x) = d(a_n, 0) < \delta$ for all $n \geq N$. Thus f_{a_n} converge pointwise to f. For the converse, suppose that f_{a_n} converge pointwise to f whenever $(a_n)_{n=0}^{\infty}$ converges to 0. By Definition 3.2.1, for every $x \in \mathbf{R}$ and $\varepsilon > 0$, there exists an $N > 0$ and $\delta > 0$ such that $d(f_{a_n}(x), f(x)) = d(f(x - a_n), f(x)) < \varepsilon$ for all $n > N$ whenever $d(x - a_n, x) = d(a_n, 0) < \delta$. Thus f is continuous.

(b) Suppose that f is uniformly continuous, by Definition 2.3.4, for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $d(f(x), f(x')) < \varepsilon$ whenever $x, x' \in \mathbf{R}$ such that $d(x, x') < \delta$. When $(a_n)_{n=0}^{\infty}$ converges to 0, for every $\delta > 0$ there is an $N > 0$ such that $d(a_n, 0) < \delta$ for all $n \geq N$. Therefore, for every $\varepsilon > 0$ there exists an $N > 0$ and $\delta > 0$ such that $d(f_{a_n}(x), f(x)) = d(f(x - a_n), f(x))) < \varepsilon$ whenever $x - a_n, x \in \mathbf{R}$ such that $d(x - a_n, x) = d(a_n, 0) < \delta$ for all $n > N$ and $x \in \mathbf{R}$. Thus f_{a_n} converge uniformly to f. Conversely, suppose that f_{a_n} converge uniformly to f. By Definition 3.2.7, for every $\varepsilon > 0$ there exists $N > 0$ and $\delta > 0$ such that $d(f_{a_n}(x), f(x)) = d(f(x - a_n), f(x)) < \varepsilon$ for every $n > N$ and $x \in \mathbf{R}$, whenever $d(a_n, 0) < \delta$ for all $n > N$. Thus by Definition 2.3.4, f is uniformly continuous. $\square$

Exercise 3.2.2.

(a) *Let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of functions from one metric space (X, d_X) to another (Y, d_Y), and let $f : X \to Y$ be another function from X to Y. Show that if $f^{(n)}$ converges uniformly to f, then $f^{(n)}$ also converges*

pointwise to f.

(b) *For each integer $n \geq 1$, let $f^{(n)} : (-1, 1) \to \mathbf{R}$ be the function $f^{(n)}(x) := x^n$. Prove that $f^{(n)}$ converges pointwise to the zero function 0, but does not converge uniformly to any function $f : (-1, 1) \to \mathbf{R}$.*

(c) *Let $g : (-1, 1) \to \mathbf{R}$ be the function $g(x) := x/(1 - x)$. With the notation as in (b), show that the partial sums $\sum_{n=1}^{N} f^{(n)}$ converges pointwise as $N \to \infty$ to g, but does not converge uniformly to g, on the open interval $(-1, 1)$. (Hint: use Lemma 7.3.3.) What would happen if we replaced the open interval $(-1, 1)$ with the closed interval $[-1, 1]$?*

3.3

Uniform convergence and continuity

Exercise 3.3.1. *Prove Theorem 3.3.1. Explain briefly why your proof requires uniform convergence, and why pointwise convergence would not suffice. (Hints: it is easiest to use the "epsilon-delta" definition of continuity from Definition 2.1.1. You may find the triangle inequality*

$$d_Y(f(x), f(x_0)) \leq d_Y(f(x), f^{(n)}(x)) + d_Y(f^{(n)}(x), f^{(n)}(x_0))$$
$$+ d_Y(f^{(n)}(x_0), f(x_0))$$

useful. Also, you may need to divide ε as $\varepsilon = \varepsilon/3 + \varepsilon/3 + \varepsilon/3$. Finally, it is possible to prove Theorem 3.3.1 from Proposition 3.3.3, but you may find it easier conceptually to prove Theorem 3.3.1 first.)

> **Theorem 3.3.1** (Uniform limits preserve continuity I). Suppose $(f^{(n)})_{n=1}^{\infty}$ is a sequence of functions from one metric space (X, d_X) to another (Y, d_Y), and suppose that this sequence converges uniformly to another function $f : X \to Y$. Let x_0 be a point in X. If the functions $f^{(n)}$ are continuous at x_0 for each n, then the limiting function f is also continuous at x_0.

Proof. Since $f^{(n)}$ converges uniformly to f, by Definition 3.2.7, for every $\varepsilon/3 > 0$ there is an $N > 0$ such that $d_Y(f(x), f^{(n)}(x)) < \varepsilon/3$ for every $n \geq N$ and $x \in X$. We also have, by Definition 2.1.1, for every $\varepsilon/3 > 0$ and n there exists a $\delta > 0$ such that $d_Y(f^{(n)}(x), f^{(n)}(x_0)) < \varepsilon/3$ whenever $d_X(x, x_0) < \delta$. Thus for every $\varepsilon > 0$ there is an $N > 0$ and $\delta > 0$ such that

$$d_Y(f(x), f(x_0)) \leq d_Y(f(x), f^{(n)}(x)) + d_Y(f^{(n)}(x), f^{(n)}(x_0))$$
$$+ d_Y(f^{(n)}(x_0), f(x_0)) < \varepsilon$$

whenever $d_X(x, x_0) < \delta$ for each $n \geq N$. Thus f is continuous at x_0. $\square$

Exercise 3.3.2. *Prove Proposition 3.3.3. (Hint: this is very similar to Theorem 3.3.1. Theorem 3.3.1 cannot be used to prove Proposition 3.3.3, however it is possible to use Proposition 3.3.3 to prove Theorem 3.3.1.)*

> **Proposition 3.3.3** (Interchange of limits and uniform limits). Let (X, d_X) and (Y, d_Y) be metric spaces, with Y complete, and let E be a subset of X. Let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of functions from E to Y, and suppose that this sequence converges uniformly in E to some function $f : E \to Y$. Let $x_0 \in X$ be an adherent point of E, and suppose that for each n the limit $\lim_{x \to x_0; x \in E} f^{(n)}(x)$ exists. Then the limit $\lim_{x \to x_0; x \in E} f(x)$ also exists, and is equal to the limit of the sequence $(\lim_{x \to x_0; x \in E} f^{(n)}(x))_{n=1}^{\infty}$; in other words we have the interchange of limits
> $$\lim_{n \to \infty} \lim_{x \to x_0; x \in E} f^{(n)}(x) = \lim_{x \to x_0; x \in E} \lim_{n \to \infty} f^{(n)}(x).$$

Proof. Since Y is complete and $(f^{(n)})_{n=1}^{\infty}$ converges to f, by Lemma 1.4.7, for every $\varepsilon > 0$ there is an $N > 0$ such that

$$d_Y(f^{(n)}(x), f^{(m)}(x)) < \varepsilon$$

for all $n, m \geq N$ and every $x \in E$. When $x \to x_0$, we have

$$d_Y(\lim_{x \to x_0; x \in E} f^{(n)}(x), \lim_{x \to x_0; x \in E} f^{(m)}(x)) < \varepsilon$$

for all $n, m \geq N$. By Definitions 1.4.6, 1.4.10, $(\lim_{x \to x_0; x \in E} f^{(n)}(x))_{n=1}^{\infty}$ converges in Y, say to L.

By triangle inequality, we have

$$d_Y(f(x), L) \le d_Y(f(x), f^{(n)}(x)) + d_Y(f^{(n)}(x), \lim_{x \to x_0; x \in E} f^{(n)}(x))$$

$$+ d_Y(\lim_{x \to x_0; x \in E} f^{(n)}(x), L).$$

From uniform convergence, for every $\varepsilon/3 > 0$ there exists an $N > 0$ such that

$$d_Y(f(x), f^{(n)}(x)) + d_Y(f^{(n)}(x) < \varepsilon/3,$$

and

$$d_Y(\lim_{x \to x_0; x \in E} f^{(n)}(x), L) < \varepsilon/3$$

for every $n \ge N$ and $x \in X$.

For such n, we can find a open ball $B(x_0, r)$ of $x \in E \cap B(x_0, r)$ such that

$$d_Y(f^{(n)}(x), \lim_{x \to x_0; x \in E} f^{(n)}(x)) < \varepsilon/3.$$

Thus we have

$$d_Y(f(x), L) < \varepsilon.$$

By Definition 3.2.1, this means that

$$\lim_{n \to \infty} \lim_{x \to x_0; x \in E} f^{(n)}(x) = \lim_{x \to x_0; x \in E} f(x) = \lim_{x \to x_0; x \in E} \lim_{n \to \infty} f^{(n)}(x),$$

as desired. $\qquad\square$

Exercise 3.3.3. *Compare Proposition 3.3.3 with Example 1.2.8. Can you now explain why the interchange of limits in Example 1.2.8 led to a false statement, whereas the interchange of limits in Proposition 3.3.3 is justified?*

Proof. $f^{(n)}(x) = x^n$ doesn't converge uniformly to f which is

$$f(x) = \begin{cases} 0 & \text{if } 0 \le x < 1 \\ 1 & \text{if } x = 1. \end{cases}$$

$\qquad\square$

Exercise 3.3.4. *Prove Proposition 3.3.4. (Hint: again, this is similar to Theorem 3.3.1 and Proposition 3.3.3, although the statements are slightly different, and one cannot deduce this directly from the other two results.)*

Proposition 3.3.4. Let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of continuous functions from one metric space (X, d_X) to another (Y, d_Y), and suppose that this sequence converges uniformly to another function $f : X \to Y$. Let $x^{(n)}$ be a sequence of points in X which converge to some limit x. Then $f^{(n)}(x^{(n)})$ converges (in Y) to $f(x)$.

Proof. Since $(f^{(n)})_{n=1}^{\infty}$ is a sequence of continuous functions, for every $\varepsilon > 0$ and n, there exists a $\delta > 0$ such that $d_Y(f^{(n)}(x), f^{(n)}(x_0)) < \varepsilon/2$ whenever $d_X(x, x_0) < \delta$. Since $(f^{(n)})_{n=1}^{\infty}$ converges uniformly to f, for every $\varepsilon > 0$ there exists an $N > 0$ such that $d_Y(f^{(n)}(x), f(x)) < \varepsilon/2$ for all $n > N$ and $x \in X$. Since $x^{(n)}$ converges to x, for every $\delta > 0$ there exists an $N > 0$ such that $d_X(x^{(n)}, x) < \delta$. Then we have $d_Y(f^{(n)}(x^{(n)}), f^{(n)}(x)) < \varepsilon/2$. Because $x \in X$, we also have $d_Y(f^{(n)}(x), f(x)) < \varepsilon/2$. Then we have following triangle inequality,

$$d_Y(f^{(n)}(x^{(n)}), f(x)) \leq d_Y(f^{(n)}(x^{(n)}), f^{(n)}(x)) + d_Y(f^{(n)}(x), f(x)) \leq \varepsilon.$$

Thus $f^{(n)}(x^{(n)})$ converges to $f(x)$. $\square$

Exercise 3.3.5. *Give an example to show that Proposition 3.3.4 fails if the phrase "converges uniformly" is replaced by "converges pointwise". (Hint: some of the examples already given earlier will already work here.)*

Proof. Consider the functions $f^{(n)} : [0, 1] \to \mathbf{R}$ defined by $f^{(n)}(x) := x^n$, and let $f : [0, 1] \to \mathbf{R}$ be the function defined by setting $f(x) := 1$ when $x = 1$ and $f(x) := 0$ when $0 \leq x < 1$. Let $x^{(n)} = 1 - \frac{1}{n}$ which converges to 1. We can see that $f(1) = 1$ but $\lim_{n \to \infty} f^{(n)}(x^{(n)}) = \lim_{n \to \infty}(1 - \frac{1}{n})^n = \frac{1}{e}$. $\square$

Exercise 3.3.6. *Prove Proposition 3.3.6. Discuss how this proposition differs from Exercise 3.2.4.*

Proposition 3.3.6 (Uniform limits preserve boundedness). Let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of functions from one metric space (X, d_X) to another (Y, d_Y), and suppose that this sequence converges uniformly to another function $f : X \to Y$. If the functions $f^{(n)}$ are bounded on X for each n, then the limiting function f is also bounded on X.

Proof. Since $f^{(n)}$ is bounded on X for each n, there is a ball $B(y_0, R_n)$ such that $f^{(n)}(x) \in B(y_0, R_n)$ for all $x \in X$ and each n. Since $(f^{(n)})_{n=1}^{\infty}$ converges uniformly to f, for every $\varepsilon > 0$ there exists an $N > 0$ such that $d_Y(f^{(n)}(x), f(x)) < \varepsilon$ for all $n > N$ and $x \in X$. Let $R = \max\{R_n : n > N\}$. Then for every $\varepsilon > 0$ there exists an $N > 0$ such that

$$d_Y(f(x), y_0) \leq d_Y(f(x), f^{(n)}(x)) + d_Y(f^{(n)}(x), y_0) \leq \varepsilon + R$$

for all $n > N$. Since ε is arbitrary, we have $d_Y(f(x), y_0) < R$. Thus f is bounded and $f(x) \in B(Y_0, R)$. $\qquad\square$

3.4 The metric of uniform convergence

Exercise 3.4.1. *Let (X, d_X) and (Y, d_Y) be metric spaces. Show that the space $B(X \to Y)$ defined in Definition 3.4.2, with the metric $d_{B(X \to Y)}$, is indeed a metric space.*

Proof. By Definition 3.4.2, we have

(a) For any $f \in B(X \to Y)$, we have
$$d_{B(X \to Y)}(f, f) = \sup_{x \in X} d_Y(f(x), f(x)) = 0.$$

(b) For any distinct $f, g \in B(X \to Y)$, we have
$$d_{B(X \to Y)}(f, g) = \sup_{x \in X} d_Y(f(x), g(x)) > 0.$$

(c) For any $f, g, h \in B(X \to Y)$, we have
$$
\begin{aligned}
d_{B(X \to Y)}(f, h) &= \sup_{x \in X} d_Y(f(x), h(x)) \\
&\leq \sup_{x \in X} d_Y(f(x), g(x)) + \sup_{x \in X} d_Y(g(x), h(x)) \\
&= d_{B(X \to Y)}(f, g) + d_{B(X \to Y)}(g, h).
\end{aligned}
$$

Thus $(B(X \to Y), d_{B(X \to Y)})$ is a metric space. $\qquad\square$

Exercise 3.4.2. *Prove Proposition 3.4.4.*

> **Proposition 3.4.4.** Let (X, d_X) and (Y, d_Y) be metric spaces. Let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of functions in $B(X \to Y)$, and let f be another function in $B(X \to Y)$. Then $(f^{(n)})_{n=1}^{\infty}$ converges to f in the metric $d_{B(X \to Y)}$ if and only if $(f^{(n)})_{n=1}^{\infty}$ converges uniformly to f.

Proof. Suppose that $(f^{(n)})_{n=1}^{\infty}$ converges to f. Then for every $\varepsilon > 0$ there exists an $N > 0$ such that

$$d_{B(X \to Y)}(f^{(n)}, f) = \sup_{x \in X} d_Y(f^{(n)}(x), f(x)) < \varepsilon$$

for all $n > N$. Then we have $d_Y(f^{(n)}(x), f(x)) < \varepsilon$ for all $n > N$ and $x \in X$. Thus $(f^{(n)})_{n=1}^{\infty}$ converges uniformly to f.

Conversely, suppose that $(f^{(n)})_{n=1}^{\infty}$ converges uniformly to f. Since $(f^{(n)})_{n=1}^{\infty}$ in $B(X \to Y)$, by Proposition 3.3.6, we have $f \in B(X \to Y)$. By Definition 3.2.7, for every $\varepsilon > 0$ there exists an $N > 0$ such that

$$d_Y(f^{(n)}(x), f(x)) < \varepsilon$$

for all $n > N$ and $x \in X$. Thus

$$d_{B(X \to Y)}(f^{(n)}, f) = \sup_{x \in X} d_Y(f^{(n)}(x), f(x)) < \varepsilon$$

and $(f^{(n)})_{n=1}^{\infty}$ converges to f in $B(X \to Y)$.   □

3.5

Series of functions; the Weierstrass M-test

Exercise 3.5.1. *Let $f^{(1)}, \cdots, f^{(N)}$ be a finite sequence of bounded functions from a metric space (X, d_X) to $\mathbf{R}$. Show that $\sum_{i=1}^{N} f^{(i)}$ is also bounded. Prove a similar claim when "bounded" is replaced by "continuous". What if "continuous" was replaced by "uniformly continuous"?*

Proof. Suppose that $f^{(1)}, \cdots, f^{(N)}$ are bounded functions, then by Exercise 3.2.4, there exist some intervals $[-R_1, R_1], \cdots, [-R_N, R_N]$ such that $f^{(1)}(x) \in [-R_1, R_1], \cdots, f^{(N)}(x) \in [-R_N, R_N]$ for all $x \in X$. Let $R := \sum_{i=1}^{N} R_i$. Then there exists a ball $[-R, R]$ such that $\sum_{i=1}^{N} f^{(i)}(x) \in [-R, R]$. Thus $\sum_{i=1}^{N} f^{(i)}$ is also bounded.

When "bounded" is replaced by "continuous", by Corollary 2.2.3, use induction we can easy to show that $\sum_{i=1}^{N} f^{(i)}$ is continuous.

When "bounded" is replaced by "uniformly continuous", by Corollary 2.2.3 and Proposition 3.4.4, $\sum_{i=1}^{N} f^{(i)}$ also is uniformly continuous. □

Exercise 3.5.2. *Prove Theorem 3.5.7. (Hint: first show that the sequence $\sum_{i=1}^{N} f^{(i)}$ is a Cauchy sequence in $C(X \to \mathbf{R})$. Then use Theorem 3.4.5.)*

> **Theorem 3.5.7** (Weierstrass M-test). Let (X, d) be a metric space, and let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of bounded real-valued continuous functions on X such that the series $\sum_{n=1}^{\infty} \|f^{(n)}\|_{\infty}$ is convergent. (Note that this is a series of plain old real numbers, not of functions.) Then the series $\sum_{n=1}^{\infty} f^{(n)}$ converges uniformly to some function f on X, and that function f is also continuous.

Proof. Since $\sum_{n=1}^{\infty} \|f^{(n)}\|_{\infty}$ is convergent, $\sum_{i=1}^{N} f^{(i)}$ is also convergent and by Lemma 1.4.7 it is a Cauchy sequence in $C(X \to \mathbf{R})$. Thus by Theorem 3.4.5, $\sum_{n=1}^{\infty} f^{(n)}$ converges to a function f in $C(X \to \mathbf{R})$. Then by Proposition 3.4.4, $\sum_{n=1}^{\infty} f^{(n)}$ converges uniformly to f. By Corollary, 3.3.2, f is continuous. □

3.6

Uniform convergence and integration

Exercise 3.6.1. *Use Theorem 3.6.1 to prove Corollary 3.6.2.*

> **Corollary 3.6.2.** Let $[a, b]$ be an interval, and let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of Riemann integrable functions on $[a, b]$ such that the series $\sum_{n=1}^{\infty} f^{(n)}$ is uniformly convergent. Then we have
> $$\sum_{n=1}^{\infty} \int_{[a,b]} f^{(n)} = \int_{[a,b]} \sum_{n=1}^{\infty} f^{(n)}.$$

Proof. Since $(f^{(n)})_{n=1}^{\infty}$ is a sequence of Riemann integrable, by the laws of Riemann integration, $\sum_{n=1}^{N} f^{(n)}$ is also Riemann integrable. Such a series

is uniformly convergent, by Theorem 3.6.1 we have

$$\lim_{n\to\infty} \int_{[a,b]} \sum_{n=1}^{N} f^{(n)} = \int_{[a,b]} \sum_{n=1}^{\infty} f^{(n)}.$$

Because integral operation is independent with sum operation, we have

$$\lim_{n\to\infty} \int_{[a,b]} \sum_{n=1}^{N} f^{(n)} = \lim_{n\to\infty} \sum_{n=1}^{N} \int_{[a,b]} f^{(n)} = \sum_{n=1}^{\infty} \int_{[a,b]} f^{(n)}.$$

Thus

$$\sum_{n=1}^{\infty} \int_{[a,b]} f^{(n)} = \int_{[a,b]} \sum_{n=1}^{\infty} f^{(n)}.$$

$\square$

Uniform convergence and derivatives

Exercise 3.7.1. *Complete the proof of Theorem 3.7.1. Compare this theorem with Example 1.2.10, and explain why this example does not contradict the theorem.*

> **Theorem 3.7.1.** Let $[a, b]$ be an interval, and for every integer $n \geq 1$, let $f_n : [a, b] \to \mathbf{R}$ be a differentiable function whose derivative $f'_n : [a, b] \to \mathbf{R}$ is continuous. Suppose that the derivatives f'_n converge uniformly to a function $g : [a, b] \to \mathbf{R}$. Suppose also that there exists a point x_0 such that the limit $\lim_{n\to\infty} f_n(x_0)$ exists. Then the functions f_n converge uniformly to a differentiable function f, and the derivative of f equals g.

Proof. Since f'_n is continuous, we see from the fundamental theorem of calculus (Theorem 11.9.4) that

$$f_n(x) - f_n(x_0) = \int_{[x_0, x]} f'_n$$

where $x \in [x_0, b]$, and

$$f_n(x) - f_n(x_0) = -\int_{[x, x_0]} f'_n$$

where $x \in [a, x_0]$.

Let L be the limit of $f_n(x_0)$ as $n \to \infty$:

$$L := \lim_{n \to \infty} f_n(x_0).$$

By hypothesis, L exists. Now, since each f_n' is continuous on $[a, b]$, and f_n' converges uniformly to g, we see by Corollary 3.3.2 that g is also continuous. Now define the function $f : [a, b] \to \mathbf{R}$ by setting

$$f(x) := L - \int_{[a, x_0]} g + \int_{[a, x]} g$$

for all $x \in [a, b]$. To finish the proof, we have to show that f_n converges uniformly to f, and that f is differentiable with derivative g.

Let $\varepsilon > 0$. Since $f_n(x_0)$ converges to L, there exists an $N > 0$ such that

$$|f_n(x_0) - L| < \varepsilon(1 + x - x_0)$$

for all $n > N$ and $x \in [x_0, b]$. By laws of Riemann integration, we also have

$$\left| \int_{[a,b]} f_n' - \int_{[a,b]} g \right| = \int_{[a,b]} |f_n' - g| < \varepsilon(b - a)$$

for all $n > N$.

To show that f_n converges uniformly to f, we divide it into two cases. When $x \in [x_0, b]$, we can see that there exists an $N > 0$ such that

$$d(f_n(x), f(x)) = \left| f_n(x_0) + \int_{[x_0, x]} f_n' - L + \int_{[a, x_0]} g - \int_{[a, x]} g \right|$$

$$\leq |f_n(x_0) - L| + \left| \int_{[x_0, x]} f_n' + 0 - \int_{[x_0, x]} g \right|$$

$$< \varepsilon + \left| \int_{[x_0, x]} f_n' - \int_{[x_0, x]} g \right|$$

$$< \varepsilon + \varepsilon(x - x_0) = \varepsilon$$

for all $n > N$ and $x \in [x_0, b]$.

A similar argument shows that $d(f_n(x), f(x)) < \varepsilon$ for all $n > N$ and $x \in [a, x_0]$. Together our conclusions, we have f_n converges uniformly to f.

Now we show that f is differentiable with derivative g. Let x' be any

point in $[a, b]$, then we have

$$\lim_{x \to x' : x \in [a,b] \backslash \{x'\}} \frac{f(x) - f(x')}{x - x'} = \lim_{x \to x' : x \in [a,b] \backslash \{x'\}} \frac{1}{x - x'} \left(\int_{[a,x]} g - \int_{[a,x']} g \right).$$

Let G to be the antiderivative of f, then we have

$$\lim_{x \to x' : x \in [a,b] \backslash \{x'\}} \frac{f(x) - f(x')}{x - x'} = \lim_{x \to x' : x \in [a,b] \backslash \{x'\}} \frac{G(x) - G(x')}{x - x'}.$$

Since G is differentiable and limit exists, hence f is also differentiable. Furthermore, the derivative of f equals to the derivative of G, which equals to g. $\square$

Exercise 3.7.2. *Prove Theorem 3.7.1 without assuming that f'_n is continuous. (This means that you cannot use the fundamental theorem of calculus. However, the mean value theorem (Corollary 10.2.9) is still available. Use this to show that if $d_\infty(f'_n, f'_m) \le \varepsilon$, then $|(f_n(x) - f_m(x)) - (f_n(x_0) - f_m(x_0))| \le \varepsilon |x - x_0|$ for all $x \in [a, b]$, and then use this to complete the proof of Theorem 3.7.1.)*

Proof. Suppose that f'_n is not continuous. Since f_n is differentiable, by Proposition 10.1.10, it is continuous. By the mean value theorem (Corollary 10.2.9), for each $x \in [a, b]$ there exists an $x' \in [x_0, x] \subseteq [a, b]$ such that

$$f'_n(x') = \frac{f_n(x) - f_n(x_0)}{x - x_0}.$$

for all n. Since f'_n converges uniformly to g, by Proposition 3.4.4, f'_n converges to g in $B([a, b] \to \mathbf{R})$. Then f'_n is a Cauchy sequence, there exists an $N > 0$ such that

$$d_\infty(f'_n, f'_m) = \sup_{x' \in [a,b]} d(f'_n(x'), f'_m(x'))$$

$$= \sup_{x' \in [a,b]} \left| \frac{(f_n(x) - f_m(x)) - (f_n(x_0) - f_m(x_0))}{x - x_0} \right| < \varepsilon$$

for all $m, n > N$. This implies that $|(f_n(x) - f_m(x)) - (f_n(x_0) - f_m(x_0))| \le \varepsilon |x - x_0|$ for all $x \in [a, b]$ and $m, n > N$.

Let L be the limit of $f_n(x_0)$ as $n \to \infty$:

$$L := \lim_{n \to \infty} f_n(x_0).$$

By hypothesis, L exists. Now define the function $f : [a, b] \to \mathbf{R}$ by setting

$$f(x) := L - \int_{[a,x_0]} g + \int_{[a,x]} g$$

for all $x \in [a, b]$. To finish the proof, we have to show that f_n converges uniformly to f, and that f is differentiable with derivative g.

Since we have

$$d_\infty(f_n, f_m) = \sup_{x \in [a,b]} |f_n(x) - f_m(x)|$$

$$\leq (f_n(x_0) - f_m(x_0)) + \varepsilon |x - x_0|$$

$$< (f_n(x_0) - f_m(x_0)) + \varepsilon(b - a).$$

For every $m, n > N$ we have $f_n(x_0) - f_m(x_0) = 0$, so that $d_\infty(f_n, f_m) < \varepsilon$. This means that f_n is a Cauchy sequence in $C([a, b] \to \mathbf{R})$. Since $C([a, b] \to \mathbf{R})$ is complete by Theorem 3.4.5, f_n converges to f. Then by Proposition 3.4.4, f_n converges uniformly to f.

Now we turn to show that f is differentiable with derivative g. Since we have

$$f_n'(x') = \frac{f_n(x) - f_n(x_0)}{x - x_0},$$

taking the limits in the both side of equality, then

$$g(x') = \frac{f(x) - f(x_0)}{x - x_0}.$$

Notice that $x' \in [x_0, x]$, if we let x converges to x_0, then x' is also converges to x_0. Thus we have

$$\lim_{x \to x_0; x \in [a,b] \setminus \{x_0\}} g(x') = \lim_{x \to x_0; x \in [a,b] \setminus \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0} = g(x_0).$$

Thus f is differentiable at x_0 and $f'(x_0) = g(x_0)$. Since this is hold for all $x \in [a, b]$, we have $f'(x) = g(x)$ and f is differentiable with derivative g. $\quad \square$

Exercise 3.7.3. *Prove Corollary 3.7.3.*

> **Corollary 3.7.3.** Let $[a, b]$ be an interval, and for every integer $n \geq 1$, let $f_n : [a, b] \to \mathbf{R}$ be a differentiable function whose derivative $f_n' : [a, b] \to \mathbf{R}$ is continuous. Suppose that the series $\sum_{n=1}^{\infty} \|f_n'\|_\infty$ is abso-

lutely convergent, where

$$\|f'_n\|_\infty := \sup_{x \in [a,b]} |f'_n(x)|$$

is the sup norm of f'_n, as defined in Definition 3.5.5. Suppose also that the series $\sum_{n=1}^\infty f_n(x_0)$ is convergent for some $x_0 \in [a,b]$. Then the series $\sum_{n=1}^\infty f_n$ converges uniformly on $[a,b]$ to a differentiable function, and in fact

$$\frac{d}{dx} \sum_{n=1}^\infty f_n(x) = \sum_{n=1}^\infty \frac{d}{dx} f_n(x)$$

for all $x \in [a,b]$.

Proof. Since $f'_n : [a,b] \to \mathbf{R}$ is continuous and $\sum_{n=1}^\infty \|f'_n\|_\infty$ is absolutely convergent, by Weierstrass M-test (Theorem 3.5.7), $\sum_{n=1}^\infty f'_n$ converges to some function $g : [a,b] \to \mathbf{R}$. Then by Theorem 3.7.1, $\sum_{n=1}^\infty f_n$ converges uniformly to a differentiable function f, and we have

$$\frac{d}{dx} \sum_{n=1}^\infty f_n(x) = \sum_{n=1}^\infty f'_n = \sum_{n=1}^\infty \frac{d}{dx} f_n(x).$$

$\square$

3.8 Uniform approximation by polynomials

Exercise 3.8.1. *Prove Lemma 3.8.5.*

Lemma 3.8.5. If $f : \mathbf{R} \to \mathbf{R}$ is continuous and supported on an interval $[a,b]$, and is also supported on another interval $[c,d]$, then $\int_{[a,b]} f = \int_{[c,d]} f$.

Proof. Suppose that f is supported on $[a,b]$ and $[c,d]$. Then we have $f(x) = 0$ for all $x \notin [a,b]$ and $x \notin [c,d]$ at the time. If $[a,b] \cap [c,d] = \emptyset$, then $\int_{[a,b]} f = 0$ from $f(x) = 0$ for all $x \notin [c,d]$. Similarly, we have $\int_{[c,d]} f = 0$. Thus $\int_{[a,b]} f = \int_{[c,d]} f = 0$. While if $[a,b] \cap [c,d] \neq \emptyset$. Then we have $\int_{[a,b]} f = \int_{[c,d]} f = \int_{[a,b] \cap [c,d]} f$ for that $f(x) = 0$ for all $x \notin [a,b] \cap [c,d]$. $\square$

Exercise 3.8.2.

(a) *Prove that for any real number $0 \leq y \leq 1$ and any natural number $n \geq 0$, that $(1 - y)^n \geq 1 - ny$. (Hint: induct on n. Alternatively, differentiate with respect to y.)*

(b) *Show that $\int_{-1}^{1}(1 - x^2)^n \, dx \geq \frac{1}{\sqrt{n}}$. (Hint: for $|x| \leq 1/\sqrt{n}$, use part (a); for $|x| \geq 1/\sqrt{n}$, just use the fact that $(1 - x^2)$ is positive. It is also possible to proceed via trigonometric substitution, but I would not recommend this unless you know what you are doing.)*

(c) *Prove Lemma 3.8.8. (Hint: choose $f(x)$ to equal $c(1 - x^2)^N$ for $x \in [-1, 1]$ and to equal zero for $x \notin [-1, 1]$, where N is a large number N, where c is chosen so that f has integral 1, and use (b).)*

> **Lemma 3.8.8** (Polynomials can approximate the identity). For every $\varepsilon > 0$ and $0 < \delta < 1$ there exists an (ε, δ)-approximation to the identity which is a polynomial P on $[-1, 1]$.

Proof.

(a) We use induction on n. This is trivial for $n = 0$ and $n = 1$. Now we suppose inductively that $(1 - y)^n \geq 1 - ny$ is hold, we need to show the case of $n + 1$. By induction hypothesis, we have
$$
\begin{aligned}
(1 - y)^{n+1} &= (1 - y)^n(1 - y) \\
&\geq (1 - ny)(1 - y) \\
&= 1 - y - ny + ny^2 \\
&\geq 1 - y - ny \\
&= 1 - (n + 1)y.
\end{aligned}
$$
This close the induction.

(b) When $|x| \leq 1/\sqrt{n}$, we have
$$
\int_{[-1,1]} (1 - x^2)^n \, dx \geq \int_{[-1,1]} (1 - nx^2) \, dx
$$
$$
= \int_{\left[\frac{-1}{\sqrt{n}}, \frac{1}{\sqrt{n}}\right]} (1 - nx^2) \, dx
$$

$$= \left[x - \frac{n}{3}x^3 \right]_{1/\sqrt{n}}^{1/\sqrt{n}}$$

$$= \frac{4}{3}\frac{1}{\sqrt{n}}.$$

Thus we have $\int_{-1}^{1}(1-x^2)^n \, dx \geq \frac{1}{\sqrt{n}}$ for $|x| \leq 1/\sqrt{n}$. When $|x| \geq 1/\sqrt{n}$, $(1-x^2)$ is positive. This implies that $\int_{-1}^{1}(1-x^2)^n \, dx$ is positive on $[-1, -1/\sqrt{n}] \cup [1/\sqrt{n}, 1]$. Thus

$$\int_{[-1,1]} (1-x^2)^n \, dx = \int_{\left[-\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}\right]} (1-x^2)^n \, dx$$

$$+ \int_{\left[-1, -\frac{1}{\sqrt{n}}\right] \cup \left[\frac{1}{\sqrt{n}}, 1\right]} (1-x^2)^n \, dx \geq \frac{1}{\sqrt{n}}.$$

(c) Define the function $f : \mathbf{R} \to \mathbf{R}$ by $f(x) := c(1-x^2)^N$ when $x \in [-1, 1]$; and $f(x) = 0$ when $x \notin [-1, 1]$. We need to show that f is an (ε, δ)-approximation to the identity. Definition 3.8.6(ab) is easy to verify. Now we use Exercise 3.8.2(b) to show Definition 3.8.6(c). Since we have

$$\int_{-\infty}^{\infty} f(x) \, dx = \int_{[-1,1]} c(1-x^2)^N \, dx = 1,$$

then

$$c = \frac{1}{\int_{[-1,1]}(1-x^2)^N \, dx}.$$

By Exercise 3.8.2(b), we have inequality $c \leq \sqrt{N}$. Then for all $\delta|x| \leq 1$, we have

$$|f(x)| \leq \sqrt{N}(1-x^2)^N \leq \varepsilon.$$

By binomial formula (Exercise 7.1.4), $f(x)$ can be written as

$$f(x) = c(1-x^2)^N = \sum_{j=0}^{N} \frac{c \cdot N!}{j!(N-j)!}x^j$$

which is a polynomial with degree N, as desired.   □

Exercise 3.8.3. *Let $f : \mathbf{R} \to \mathbf{R}$ be a compactly supported, continuous function. Show that f is bounded and uniformly continuous. (Hint: the*

idea is to use Proposition 2.3.2 and Theorem 2.3.5, but one must first deal with the issue that the domain $\mathbf{R}$ of f is non-compact.)

Proof. Since f is compactly supported, by Definition 3.8.4, there is some interval $[a, b]$ such that $f(x) = 0$ for all $x \notin [a, b]$. Thus for $x \notin [a, b]$, by definition of f and Definition 2.3.4, f is bounded and uniformly continuous. Now we consider the function $f|_{[a,b]} : [a, b] \to \mathbf{R}$. By Heine-Borel theorem (Theorem 1.5.7) $[a, b]$ is compact. Then by Proposition 2.3.2 and Theorem 2.3.5, f is bounded and uniformly continuous on $[a, b]$. Thus f is bounded and uniformly continuous on $\mathbf{R}$. $\qquad\square$

Exercise 3.8.4. *Prove Proposition 3.8.11. (Hint: to show that $f * g$ is continuous, use Exercise 3.8.3.)*

> **Proposition 3.8.11** (Basic properties of convolution). Let $f : \mathbf{R} \to \mathbf{R}$, $g : \mathbf{R} \to \mathbf{R}$, and $h : \mathbf{R} \to \mathbf{R}$ be continuous, compactly supported functions. Then the following statements are true.
>
> (a) The convolution $f * g$ is also a continuous, compactly supported function.
>
> (b) (Convolution is commutative) We have $f * g = g * f$; in other words
> $$f * g(x) = \int_{-\infty}^{\infty} f(y) g(x - y) \, dy$$
> $$= \int_{-\infty}^{\infty} g(y) f(x - y) \, dy$$
> $$= g * f(x).$$
>
> (c) (Convolution is linear) We have $f * (g + h) = f * g + f * h$. Also, for any real number c, we have $f * (cg) = (cf) * g = c(f * g)$.

Proof. (a) First, we show that $f * g$ is continuous. Since f and g are continuous and compactly supported, by Exercise 3.8.3, f and g is bounded and uniformly continuous. Then there is a real number M such that $|f(x)| \leq M$

for all x. Let $\varepsilon > 0$, there is a $\delta > 0$ and $b > a$ such that
$$|g(x) - g(x_0)| < \frac{\varepsilon}{M(b-a)}$$
whenever $|x - x_0| < \delta$. Now we use boundedness of f and continuity of g to prove that $f * g$ is continuous. For every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$
\begin{aligned}
|f * g(x) - f * g(x_0)| &= \left| \int_{-\infty}^{\infty} f(y)g(x-y)\,dy - \int_{-\infty}^{\infty} f(y)g(x_0-y)\,dy \right| \\
&= \int_{-\infty}^{\infty} |f(y)g(x-y) - f(y)g(x_0-y)|\,dy \\
&\leq \int_{-\infty}^{\infty} |f(y)g(x-y) - f(y)g(x_0-y)|\,dy \\
&= \int_{-\infty}^{\infty} |f(y)| \cdot |g(x-y) - g(x_0-y)|\,dy \\
&\leq \int_{-\infty}^{\infty} \frac{\varepsilon}{M(b-a)} |f(y)|\,dy
\end{aligned}
$$

whenever $|x - x_0| < \delta$. Since f is compactly supported, it is supported on some interval $[a, b]$. Then we have

$$|f * g(x) - f * g(x_0)| \leq \frac{\varepsilon}{M(b-a)} \int_{[a,b]} |f(x)|\,dx = \varepsilon.$$

Thus $f * g$ is continuous.

Now we show that $f * g$ is compactly supported. Suppose that f is supported on $[a, b]$, and g is supported on $[c, d]$. Since f is supported on $[a, b]$, we have

$$f * g(x) = \int_{-\infty}^{\infty} f(y)g(x-y)\,dy = \int_{[a,b]} f(y)g(x-y)\,dy.$$

Since g is supported on $[c, d]$, if $x - y \notin [c, d]$, i.e., $x \notin [a + c, b + d]$, we will have $f * g(x) = 0$. Thus $f * g$ is supported on $[a + c, b + d]$ and is compactly supported.

(b) Let $z = x - y$, then

$$
\begin{aligned}
f * g(x) &= \int_{-\infty}^{\infty} f(y)g(x-y)\,dy \\
&= \int_{[a,b]} f(y)g(x-y)\,dy
\end{aligned}
$$

$$= -\int_{[d,c]} g(z) f(x-z)\, dz$$

$$= \int_{[c,d]} g(z) f(x-z)\, dz$$

$$= \int_{[c,d]} g(y) f(x-y)\, dy$$

$$= \int_{-\infty}^{\infty} g(y) f(x-y)\, dy$$

$$= g * f(x).$$

Notice that second equation is hold for f is supported on $[a, b]$; third equation we use $z = x-y$, and $a \le y = a-x \le b$ implies $d \le x-b \le z \le x-a \le c$ (use (a) that $f * g$ is supported on $a+c \le x \le b+d$). Remaining equations are trivial and we have $f * g(x) = g * f(x)$.

(c) By definition, we have

$$f * (g + h)(x) = \int_{-\infty}^{\infty} f(y)(g+h)(x-y)\, dy$$

$$= \int_{-\infty}^{\infty} f(y)g(x-y) + f(y)h(x-y)\, dy$$

$$= \int_{-\infty}^{\infty} f(y)g(x-y)\, dy + \int_{-\infty}^{\infty} f(y)h(x-y)\, dy$$

$$= f * g(x) + f * h(x).$$

For any real number c,

$$f * (cg)(x) = \int_{-\infty}^{\infty} f(y)(cg)(x-y)\, dy$$

$$= \int_{-\infty}^{\infty} cf(y)g(x-y)\, dy$$

$$= \int_{-\infty}^{\infty} (cf)(y)g(x-y)\, dy$$

$$= (cf) * g(x),$$

and

$$f * (cg)(x) = \int_{-\infty}^{\infty} f(y)(cg)(x-y)\, dy$$

$$= \int_{-\infty}^{\infty} cf(y)g(x-y)\,dy$$

$$= c \int_{-\infty}^{\infty} f(y)g(x-y)\,dy$$

$$= c(f * g(x)).$$

Thus $f * (cg) = (cf) * g = c(f * g)$. $\square$

Exercise 3.8.5. *Let $f : \mathbf{R} \to \mathbf{R}$ and $g : \mathbf{R} \to \mathbf{R}$ be continuous, compactly supported functions. Suppose that f is supported on the interval $[0,1]$, and g is constant on the interval $[0,2]$ (i.e., there is a real number c such that $g(x) = c$ for all $x \in [0,2]$). Show that the convolution $f * g$ is constant on the interval $[1,2]$.*

Proof. For all $x \in [0,1]$, we have

$$f * g(x) = \int_{-\infty}^{\infty} f(y)g(x-y)\,dy = \int_{[0,1]} f(y)g(x-y)\,dy$$

since f is supported on $[0,1]$. Since $g(x) = c$ for all $x \in [0,2]$, we have

$$f * g(x) = c \int_{[0,1]} f(y)\,dy$$

for all $x - y \in [0,2]$, i.e., $x \in [1,2]$. Since $\int_{[0,1]} f(y)\,dy$ equals to some real number, thus $f * g$ is constant on the interval $[1,2]$. $\square$

Exercise 3.8.6.

(a) *Let g be an (ε, δ)-approximation to the identity. Show that $1 - 2\varepsilon \le \int_{[-\delta,\delta]} g \le 1$.*

(b) *Prove Lemma 3.8.14. (Hint: begin with the identity*

$$f * g(x) = \int f(x-y)g(y)\,dy = \int_{[-\delta,\delta]} f(x-y)g(y)\,dy$$

$$+ \int_{[\delta,1]} f(x-y)g(y)\,dy + \int_{[-1,-\delta]} f(x-y)g(y)\,dy.$$

The idea is to show that the first integral is close to $f(x)$, and that the second and third integrals are very small. To achieve the former task,

use (a) and the fact that $f(x)$ and $f(x - y)$ are within ε of each other;
to achieve the latter task, use property (c) of the approximation to the
identity and the fact that f is bounded.)

Lemma 3.8.14. Let $f : \mathbf{R} \to \mathbf{R}$ be a continuous function supported on
$[0, 1]$, which is bounded by some $M > 0$ (i.e., $|f(x)| \leq M$ for all $x \in \mathbf{R}$),
and let $\varepsilon > 0$ and $0 < \delta < 1$ be such that one has $|f(x) - f(y)| < \varepsilon$
whenever $x, y \in \mathbf{R}$ and $|x - y| < \delta$. Let g be any (ε, δ)-approximation to
the identity. Then we have

$$|f * g(x) - f(x)| \leq (1 + 4M)\varepsilon$$

for all $x \in [0, 1]$.

Proof.

(a) Let $\varepsilon > 0$ and $0 < \delta < 1$. Let g be an (ε, δ)-approximation to the
    identity. Since $[-\delta, \delta] \subseteq [-\infty, \infty]$, by Definition 3.8.6(b) we have
    $\int_{[-\delta,\delta]} g \leq 1$. Since $|g(x)| \leq \varepsilon$ for all $\delta \leq |x| \leq 1$, we have

    $$\int_{[-1,-\delta]\cup[\delta,1]} g \leq \int_{[-1,-\delta]\cup[\delta,1]} \varepsilon = 2(1 - \delta)\varepsilon.$$

    By Definition 3.8.6, we have

    $$\int_{[-\delta,\delta]} g = \int_{[-1,1]} g - \int_{[-1,-\delta]\cup[\delta,1]} g \geq 1 - 2(1 - \delta)\varepsilon \geq 1 - 2\varepsilon$$

    as desired.

(b) By Definition 3.8.6(a), we have

    $$f * g(x) = \int_{[-1,1]} f(x - y)g(y)\, dy = \int_{[-\delta,\delta]} f(x - y)g(y)\, dy$$

    $$+ \int_{[\delta,1]} f(x - y)g(y)\, dy + \int_{[-1,-\delta]} f(x - y)g(y)\, dy.$$

    Use the fact that $\int_{[-1,1]} g(y)\, dy = 1$, we have

    $$f(x) = f(x) \int_{[-1,1]} g(y)\, dy = \int_{[-\delta,\delta]} f(x)g(y)\, dy$$

    $$+ \int_{[\delta,1]} f(x)g(y)\, dy + \int_{[-1,-\delta]} f(x)g(y)\, dy.$$

Then $f * g(x) - f(x)$ can be divided into three parts. For the first part, we have

$$\left| \int_{[-\delta,\delta]} f(x-y)g(y)\, dy - \int_{[-\delta,\delta]} f(x)g(y)\, dy \right|$$

$$= \int_{[-\delta,\delta]} |f(x-y) - f(x)| \cdot |g(y)|\, dy$$

$$\leq \varepsilon \int_{[-\delta,\delta]} |g(y)| \leq \varepsilon,$$

whenever $|x - y - x| = |y| \leq \delta$. For the second part, we have

$$\left| \int_{[\delta,1]} f(x-y)g(y)\, dy - \int_{[\delta,1]} f(x)g(y)\, dy \right|$$

$$\leq \int_{[\delta,1]} |f(x-y)g(y)|\, dy + \int_{[\delta,1]} |f(x)g(y)|\, dy$$

$$\leq 2M\varepsilon(1-\delta) \leq 2M\varepsilon.$$

This is similar to show the third part that

$$\left| \int_{[-1,-\delta]} f(x-y)g(y)\, dy - \int_{[-1,-\delta]} f(x)g(y)\, dy \right| \leq 2M\varepsilon.$$

Then we can easy to see that $|f * g(x) - f(x)|$ less than the sum of above three parts, and we have

$$|f * g(x) - f(x)| \leq \varepsilon + 2M\varepsilon + 2M\varepsilon = (1 + 4M)\varepsilon$$

as desired.   □

Exercise 3.8.7. *Prove Corollary 3.8.15. (Hint: combine Exercise 3.8.3, Lemma 3.8.8, Lemma 3.8.13, and Lemma 3.8.14.)*

> **Corollary 3.8.15** (Weierstrass approximation theorem I). Let $f : \mathbf{R} \to \mathbf{R}$ be a continuous function supported on $[0, 1]$. Then for every $\varepsilon > 0$, there exists a function $P : \mathbf{R} \to \mathbf{R}$ which is polynomial on $[0, 1]$ and such that $|P(x) - f(x)| \leq \varepsilon$ for all $x \in [0, 1]$.

Proof. Let $\varepsilon > 0$. By Lemma 3.8.8, for every $\varepsilon_0 > 0$ and $0 < \delta < 1$ there is a polynomial g on $[-1, 1]$ which is an (ε_0, δ)-approximation to the identity. Then by Lemma 3.8.13, $f * g$ is a polynomial on $[0, 1]$. Let $P := f * g$. Since

f is continuous function supported on $[0, 1]$, by Exercise 3.8.3, f is bounded and uniformly continuous. Then by Lemma 3.8.14, for $\varepsilon_0 = \frac{\varepsilon}{1+4M}$ we have

$$|f * g(x) - f(x)| = |P(x) - f(x)| \leq (1 + 4M)\varepsilon_0 = \varepsilon$$

for all $x \in [0, 1]$. □

Exercise 3.8.8. *Let $f : [0, 1] \to \mathbf{R}$ be a continuous function, and suppose that $\int_{[0,1]} f(x)x^n \, dx = 0$ for all non-negative integers $n = 0, 1, 2, \cdots$. Show that f must be the zero function $f \equiv 0$. (Hint: first show that $\int_{[0,1]} f(x)P(x) \, dx = 0$ for all polynomials P. Then, using the Weierstrass approximation theorem, show that $\int_{[0,1]} f(x)f(x) \, dx = 0$.)*

Proof. Since $\int_{[0,1]} f(x)x^n \, dx = 0$ for all non-negative integers $n = 0, 1, 2, \cdots$, we have

$$\int_{[0,1]} f(x)P(x) \, dx = \sum_{j=0}^{n} c_j \int_{[0,1]} f(x)x^j \, dx = 0$$

for all polynomials P. Then by Weierstrass approximation theorem (Corollary 3.8.15), there is a polynomial P on $[0, 1]$ such that

$$\int_{[0,1]} f(x)P(x) \, dx = \int_{[0,1]} f(x)f(x) = 0.$$

This implies that $f(x) = 0$ for all $x \in [0, 1]$. □

Exercise 3.8.9. *Prove Lemma 3.8.16.*

> **Lemma 3.8.16.** Let $f : [0, 1] \to \mathbf{R}$ be a continuous function which equals 0 on the boundary of $[0, 1]$, i.e., $f(0) = f(1) = 0$. Let $F : \mathbf{R} \to \mathbf{R}$ be the function defined by setting $F(x) := f(x)$ for $x \in [0, 1]$ and $F(x) := 0$ for $x \notin [0, 1]$. Then F is also continuous.

Proof. Since f is continuous on $[0, 1]$, then F is also continuous on $[0, 1]$. For $x \notin [0, 1]$, constant function is obviously continuous. Thus F is continuous on $x \notin [0, 1]$. Now we require that $\lim_{x \to 1-} F(x) = \lim_{x \to 1+} F(x)$ and $\lim_{x \to 0-} F(x) = \lim_{x \to 0+} F(x)$. Since $f(0) = f(1) = 0$, we have

$$\lim_{x \to 1-} F(x) = \lim_{x \to 0+} F(x) = 0.$$

By definition of F we have

$$\lim_{x \to 1-} F(x) = \lim_{x \to 0-} F(x) = 0.$$

Thus F is continuous on $\mathbf{R}$ as desired. $\square$

CHAPTER IV

Power series

4.1

Formal power series

Exercise 4.1.1. *Prove Theorem 4.1.6. (Hints: for (a) and (b), use the root test (Theorem 7.5.1). For (c), use the Weierstrass M-test (Theorem 3.5.7). For (d), use Theorem 3.7.1. For (e), use Corollary 3.6.2.*

> **Theorem 4.1.6.** Let $\sum_{n=0}^{\infty} c_n(x-a)^n$ be a formal power series, and let R be its radius of convergence.
>
> (a) (Divergence outside of the radius of convergence) If $x \in \mathbf{R}$ is such that $|x - a| > R$, then the series $\sum_{n=0}^{\infty} c_n(x-a)^n$ is divergent for that value of x.
>
> (b) (Convergence inside the radius of convergence) If $x \in \mathbf{R}$ is such that $|x-a| < R$, then the series $\sum_{n=0}^{\infty} c_n(x-a)^n$ is absolutely convergent for that value of x.
>
> For parts (c)-(e) we assume that $R > 0$ (i.e., the series converges at at least one other point than $x = a$). Let $f : (a - R, a + R) \to \mathbf{R}$ be the

function $f(x) := \sum_{n=0}^{\infty} c_n(x-a)^n$; this function is guaranteed to exist by (b).

(c) (Uniform convergence on compact sets) For any $0 < r < R$, the series $\sum_{n=0}^{\infty} c_n(x-a)^n$ converges uniformly to f on the compact interval $[a-r, a+r]$. In particular, f is continuous on $(a-R, a+R)$.

(d) (Differentiation of power series) The function f is differentiable on $(a-R, a+R)$, and for any $0 < r < R$, the series $\sum_{n=0}^{\infty} nc_n(x-a)^{n-1}$ converges uniformly to f' on the interval $[a-r, a+r]$.

(e) (Integration of power series) For any closed interval $[y, z]$ contained in $(a-R, a+R)$, we have

$$\int_{[y,z]} f = \sum_{n=0}^{\infty} c_n \frac{(z-a)^{n+1} - (y-a)^{n+1}}{n+1}.$$

Proof.

(a) If $|x-a| > R$, then we have $|c_n(x-a)^n| > c_n R^n$. Since we have

$$\limsup_{n\to\infty} |c_n R^n|^{1/n} = \limsup_{n\to\infty} \frac{|c_n|^{1/n}}{\limsup_{n\to\infty} |c_n|^{1/n}} = 1,$$

this implies that $\alpha = \limsup_{n\to\infty} |c_n(x-a)^n|^{1/n} > 1$. By the root test (Theorem 7.5.1), $\sum_{n=0}^{\infty} c_n(x-a)^n$ is divergent.

(b) If $|x-a| < R$, then we have $|c_n(x-a)^n| < c_n R^n$. Since we have $\limsup_{n\to\infty} |c_n R^n|^{1/n} = 1$, this implies that $\alpha = \limsup_{n\to\infty} |c_n(x-a)^n|^{1/n} < 1$. By the root test, $\sum_{n=0}^{\infty} c_n(x-a)^n$ is absolutely convergent.

(c) Since c_n is finite for all n and $|x-a| < R$, we have $|(x-a)^n| < R^n$. Then there is a real number $M > 0$ such that $c_n(x-a)^n < M$ for all n. From (b), we have $\sum_{n=0}^{\infty} c_n(x-a)^n$ is absolutely convergent for all $x \in [a-r, a+r]$. Then by the Weierstrass M-test, $\sum_{n=0}^{\infty} c_n(x-a)^n$ converges uniformly to f on $[a-r, a+r]$. Since $\sum_{n=0}^{\infty} c_n(x-a)^n$ converges on $(a-R, a+R)$, f is continuous on $(a-R, a+R)$.

(d) For any $0 < r < R$ and $n \geq 1$, let $f_n : [a-r, a+r] \to \mathbf{R}$ to be

the function $f_n(x) := \sum_{k=0}^{n} c_k (x - a)^k$. Then the derivative of f_n is $f_n' : [a - r, a + r] \to \mathbf{R}$ that $f_n'(x) = \sum_{k=1}^{n} k c_k (x - a)^{k-1}$. By Theorem 4.1.1(b), there is a function $g : [a-r, a+r] \to \mathbf{R}$ such that f_n' converges uniformly to $g(x) := \sum_{n=1}^{\infty} n c_n (x - a)^{n-1}$. Since $\sum_{n=0}^{\infty} c_n (x - a)^n$ is convergent on $[a-r, a+r]$, thus $\lim_{n \to \infty} f(x_0)$ exists. Then by Theorem 3.7.1, f_n converges uniformly to differentiable function f, this is easy to see that $f(x) = \sum_{n=0}^{\infty} c_n (x - a)^n$; and the derivative of f equals to g.

(e) By Corollary 3.6.2, we have

$$
\begin{aligned}
\int_{[y,z]} f &= \int_{[y,z]} \sum_{n=0}^{\infty} c_n (x - a)^n \\
&= \sum_{n=0}^{\infty} \int_{[y,z]} c_n (x - a)^n \\
&= \sum_{n=0}^{\infty} \left[\frac{c_n (x - a)^{n+1}}{n + 1} \right]_y^z \\
&= \sum_{n=0}^{\infty} c_n \frac{(z - a)^{n+1} - (y - a)^{n+1}}{n + 1}.
\end{aligned}
$$

$\square$

4.2

Real analytic function

Exercise 4.2.1. *Let $n \geq 0$ be an integer, let c, a be real numbers, and let f be the function $f(x) := c(x-a)^n$. Show that f is infinitely differentiable, and that $f^{(k)}(x) = c\frac{n!}{(n-k)!}(x-a)^{n-k}$ for all integers $0 \leq k \leq n$. What happens when $k > n$?*

Proof. We use induction on k. When $k = 1$, we can know easily that f is differentiable and $f'(x) = cn(x-a)^{n-1}$. Now we suppose that f is $k-1$ times differentiable, and we have

$$f^{(k-1)}(x) = c\frac{n!}{(n-k+1)!}(x-a)^{n-k+1}.$$

Then

$$f^{(k)} = \left(f^{(k-1)}\right)'(x) = c\frac{n!}{(n-k)!}(x-a)^{n-k}$$

as desired. $\square$

Exercise 4.2.2. *Show that the function f defined in Example 4.2.2 is real analytic on all of $\mathbf{R} \setminus \{1\}$.*

Proof. Consider the function $f : \mathbf{R} \setminus \{1\} \to \mathbf{R}$ defined by $f(x) := 1/(1-x)$. Consider the Taylor series

$$\sum_{n=0}^{\infty} \frac{1}{(1-a)^{n+1}}(x-a)^n.$$

Then by Lemma 7.3.3, we have $|(x-a)/(1-a)| < 1$ for all $x \in (a-r, a+r)$ where $r = |1-a|$, and

$$\sum_{n=0}^{\infty} \frac{1}{(1-a)^{n+1}}(x-a)^n = \frac{1}{1-a}\sum_{n=0}^{\infty}\frac{1}{(1-a)^n}(x-a)^n$$

$$= \frac{1}{1-a} \cdot \frac{1}{1-\frac{x-a}{1-a}}$$

$$= \frac{1}{1-x}.$$

Thus the series above converges to f and f is analytic function at a. $\qquad\square$

Exercise 4.2.3. *Prove Proposition 4.2.6. (Hint: induct on k and use Theorem 4.1.6(d)).*

> **Proposition 4.2.6** (Real analytic functions are k-times differentiable).
> Let E be a subset of $\mathbf{R}$, let a be an interior point of E, and let f be a function which is real analytic at a, thus there is an $r > 0$ for which we have the power series expansion
> $$f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$$
> for all $x \in (a-r, a+r)$. Then for every $k \geq 0$, the function f is k-times differentiable on $(a-r, a+r)$, and for each $k \geq 0$ the kth derivative is given by
> $$f^{(k)}(x) = \sum_{n=0}^{\infty} c_{n+k}(n+1)(n+2)\cdots(n+k)(x-a)^n$$
> $$= \sum_{n=0}^{\infty} c_{n+k}\frac{(n+k)!}{n!}(x-a)^n$$
> for all $x \in (a-r, a+r)$.

Proof. We use induction on k. This is trivial for $k = 0$. When $k = 1$, by Proposition 4.1.6(d), f is once differentiable on $(a-r, a+r)$ and we have
$$f'(x) = \sum_{n=1}^{\infty} nc_n(x-a)^{n-1}$$
$$= \sum_{n=0}^{\infty} c_{n+1}(n+1)(x-a)^n.$$
Now we suppose inductively that f is $k-1$-times differentiable and we have
$$f^{(k-1)}(x) = \sum_{n=0}^{\infty} c_{n+k-1}\frac{(n+k-1)!}{n!}(x-a)^n.$$
By Proposition 4.1.6(d) again, $f^{(k-1)}$ is differentiable and we have
$$\left(f^{(k-1)}(x)\right)' = f^{(k)}(x) = \sum_{n=1}^{\infty} c_{n+k-1}\frac{(n+k-1)!}{(n-1)!}(x-a)^{n-1}$$

$$= \sum_{n=0}^{\infty} c_{n+k} \frac{(n+k)!}{n!} (x-a)^n$$

for all $x \in (a-r, a+r)$, as desired. □

Exercise 4.2.4. *Use Proposition 4.2.6 and Exercise 4.2.1 to prove Corollary 4.2.10.*

> **Corollary 4.2.10** (Taylor's formula). Let E be a subset of $\mathbf{R}$, let a be an interior point of E, and let $f : E \to \mathbf{R}$ be a function which is real analytic at a and has the power series expansion
>
> $$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$$
>
> for all $x \in (a-r, a+r)$ and some $r > 0$. Then for any integer $k \geq 0$, we have
>
> $$f^{(k)}(a) = k! c_k,$$
>
> where $k! := 1 \times 2 \times \cdots \times k$ (and we adopt the convention that $0! = 1$). In particular, we have *Taylor's formula*
>
> $$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n$$
>
> for all x in $(a-r, a+r)$.

Proof. By Proposition 4.2.6 we have

$$f^{(k)} = \sum_{n=0}^{\infty} c_{n+k} \frac{(n+k)!}{n!} (x-a)^n$$

for all $k \geq 0$ and $x \in (a-r, a+r)$. Since for all $n > 0$ parts equal to zero when $x = a$, we have

$$f^{(k)}(a) = k! c_k.$$

This implies that $c_n = f^{(n)}(a)/n!$, taking this into $f(x)$, we have

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n$$

as desired. □

4.3

Abel's theorem

Exercise 4.3.1. *Prove Lemma 4.3.2. (Hint: first work out the relationship between the partial sums $\sum_{n=0}^{N}(a_{n+1} - a_n)b_n$ and $\sum_{n=0}^{N} a_{n+1}(b_{n+1} - b_n)$.)*

> **Lemma 4.3.2** (Summation by parts formula). Let $(a_n)_{n=0}^{\infty}$ and $(b_n)_{n=0}^{\infty}$ be sequences of real numbers which converge to limits A and B respectively, i.e., $\lim_{n\to\infty} a_n = A$ and $\lim_{n\to\infty} b_n = B$. Suppose that the sum $\sum_{n=0}^{\infty}(a_{n+1} - a_n)b_n$ is convergent. Then the sum $\sum_{n=0}^{\infty} a_{n+1}(b_{n+1} - b_n)$ is also convergent, and
> $$\sum_{n=0}^{\infty}(a_{n+1} - a_n)b_n = AB - a_0b_0 - \sum_{n=0}^{\infty} a_{n+1}(b_{n+1} - b_n).$$

Proof. Since we have

$$\sum_{n=0}^{N}(a_{n+1} - a_n)b_n = \sum_{n=0}^{N} a_{n+1}b_n - \sum_{n=0}^{N} a_n b_n$$
$$= \sum_{n=0}^{N} a_{n+1}b_n - \sum_{n=0}^{N} a_{n+1}b_{n+1} - a_0b_0 + a_{N+1}b_{N+1}$$
$$= a_{N+1}b_{N+1} - a_0b_0 - \sum_{n=0}^{N} a_{n+1}(b_{n+1} - b_n).$$

Since $\lim_{n\to\infty} a_n = A$, $\lim_{n\to\infty} b_n = B$ and $\sum_{n=0}^{\infty}(a_{n+1} - a_n)b_n$ is convergent, when we take N to infinity, left side of above equation must convergent. This implies that $\sum_{n=0}^{\infty} a_{n+1}(b_{n+1} - b_n)$ is convergent, and

$$\sum_{n=0}^{\infty}(a_{n+1} - a_n)b_n = AB - a_0b_0 - \sum_{n=0}^{\infty} a_{n+1}(b_{n+1} - b_n)$$

as desired. $\square$

Multiplication of power series

<div style="text-align: center">THERE IS NO EXERCISES IN SECTION 4.4.</div>

The exponential and logarithm functions

Exercise 4.5.1. *Prove Theorem 4.5.2. (Hints: for part (a), use the ratio test. For parts (bc), use Theorem 4.1.6. For part (d), use Theorem 4.4.1. For part (e), use part (d). For part (f), use part (d), and prove that* $\exp(x) > 1$ *when x is positive. You may find the binomial formula from Exercise 7.1.4 to be useful.*

> **Theorem 4.5.2** (Basic properties of exponential).
>
> (a) For every real number x, the series $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ is absolutely convergent. In particular, $\exp(x)$ exists and is real for every $x \in \mathbf{R}$, the power series $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ has an infinite radius of convergence, and exp is a real analytic function on $(-\infty, \infty)$.
>
> (b) exp is differentiable on $\mathbf{R}$, and for every $x \in \mathbf{R}$, $\exp'(x) = \exp(x)$.
>
> (c) exp is continuous on $\mathbf{R}$, and for every interval $[a, b]$, we have $\int_{[a,b]} \exp(x)\,dx = \exp(b) - \exp(a)$.
>
> (d) For every $x, y \in \mathbf{R}$, we have $\exp(x + y) = \exp(x)\exp(y)$.
>
> (e) We have $\exp(0) = 1$. Also, for every $x \in \mathbf{R}$, $\exp(x)$ is positive, and $\exp(-x) = 1/\exp(x)$.
>
> (f) exp is strictly monotone increasing: in other words, if x, y are real numbers, then we have $\exp(y) > \exp(x)$ if and only if $y > x$.

Proof.

(a) Use the ratio test (Theorem 7.5.1), for $\alpha = \limsup_{n\to\infty} |x^n/n!|^{1/n} = \limsup_{n\to\infty} |x|/n! = 0$ for all $x \in \mathbf{R}$. Thus $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ is absolutely convergent.

(b) By Theorem 4.1.6(d), $\exp(x)$ is differentiable on $\mathbf{R}$, and for every $x \in \mathbf{R}$, we have

$$\exp'(x) = \frac{d}{dx}\sum_{n=0}^{\infty}\frac{x^n}{n!} = \sum_{n=0}^{\infty}\frac{nx^{n-1}}{n!} = \sum_{n=1}^{\infty}\frac{x^{n-1}}{(n-1)!} = \exp(x).$$

(c) exp is continuous on $\mathbf{R}$ from (b). By Theorem 4.1.6(d), we have

$$
\begin{aligned}
\int_{[a,b]} \exp(x)\,dx &= \sum_{n=0}^{\infty}\frac{b^{n+1}-a^{n+1}}{n!(n+1)} \\
&= \sum_{n=0}^{\infty}\frac{b^{n+1}-a^{n+1}}{(n+1)!} \\
&= \sum_{n=0}^{\infty}\frac{b^{n+1}}{(n+1)!} - \sum_{n=0}^{\infty}\frac{a^{n+1}}{(n+1)!} \\
&= \sum_{n=0}^{\infty}\frac{b^n}{n!} - \sum_{n=0}^{\infty}\frac{a^n}{n!} \\
&= \exp(b) - \exp(a)
\end{aligned}
$$

as desired.

(d) By Definition 4.5.1 we have (see Exercise 7.1.4)

$$\exp(x+y) = \sum_{n=0}^{\infty}\frac{(x+y)^n}{n!} = \sum_{n=0}^{\infty}\sum_{j=0}^{n}\frac{x^{n-j}y^j}{j!(n-j)!}$$

By Theorem 4.1.1 we have

$$\exp(x)\exp(y) = \left(\sum_{n=0}^{\infty}\frac{x^n}{n!}\right)\left(\sum_{n=0}^{\infty}\frac{y^n}{n!}\right) = \sum_{n=0}^{\infty}\frac{x^{n-j}y^j}{j!(n-j)!}.$$

Thus $\exp(x+y) = \exp(x)\exp(y)$.

(e) This is easy to verify that $\exp(0) = 1$ by Definition 4.5.1. Now we use (d) to show that $\exp(x)$ is positive for all $x \in \mathbf{R}$ and $\exp(-x) = 1/\exp(x)$. Since for every $x \in \mathbf{R}$, we have $x = x/2 + x/2$, then by (b)

$$\exp(x) = \exp(x/2 + x/2) = (\exp(x/2))^2.$$

$\exp(x/2)$ is real for every $x \in \mathbf{R}$, thus $\exp(x)$ cannot be negative. Then by definition of exp, it also cannot be zero. Thus $\exp(x)$ is positive for every $x \in \mathbf{R}$. For the second assertion, since

$$\exp(0) = \exp(x + (-x)) = \exp(x)\exp(-x)$$

and $\exp(0) = 1$, then we have $\exp(-x) = 1/\exp(x)$, as desired.

(*f*) Since $exp'(x) = exp(x)$ and $\exp(x) > 0$ for every $x \in \mathbf{R}$, Suppose that $\exp(y) > \exp(x)$, then we have

$$\exp(y)/\exp(x) = \exp(y)\exp(-x) = \exp(y - x) > 1$$

CHAPTER VI

Several variable differential calculus

6.1
Linear transformations

Exercise 6.1.1. *Prove Lemma 6.1.2.*

> **Lemma 6.1.2 ($\mathbf{R}^n$ is a vector space).** Let x, y, z be vectors in $\mathbf{R}^n$, and let c, d be real numbers. Then we have the *commutativity property* $x + y = y + x$, the *additive associativity property* $(x + y) + z = x + (y + z)$, the *additive identity property* $x + 0 = 0 + x = x$, the *additive inverse property* $x + (-x) = (-x) + x = 0$, the *multiplicative associativity property* $(cd)x = c(dx)$, the *distributivity properties* $c(x + y) = cx + cy$ and $(c + d)x = cx + dx$, and the *multiplicative identity property* $1x = x$.

Proof.

(a) Commutativity property.

$$
\begin{aligned}
x + y &= (x_i)_{1 \leq i \leq n} + (y_i)_{1 \leq i \leq n} \\
&= (x_i + y_i)_{1 \leq i \leq n} \\
&= (y_i + x_i)_{1 \leq i \leq n}
\end{aligned}
$$

$$= (y_i)_{1 \le i \le n} + (x_i)_{1 \le i \le n}$$
$$= y + x.$$

(b) Additive associativity property.

$$(x + y) + z = ((x_i)_{1 \le i \le n} + (y_i)_{1 \le i \le n}) + (z_i)_{1 \le i \le n}$$
$$= (x_i + y_i)_{1 \le i \le n} + (z_i)_{1 \le i \le n}$$
$$= ((x_i + y_i) + z_i)_{1 \le i \le n}$$
$$= (x_i + (y_i + z_i))_{1 \le i \le n}$$
$$= (x_i)_{1 \le i \le n} + (y_i + z_i)_{1 \le i \le n}$$
$$= (x_i)_{1 \le i \le n} + ((y_i)_{1 \le i \le n} + (z_i)_{1 \le i \le n})$$
$$= x + (y + z).$$

(c) Additive identity property.

$$x + 0 = (x_i)_{1 \le i \le n} + 0_{\mathbf{R}^n}$$
$$= (x_i + 0)_{1 \le i \le n}$$
$$= (0 + x_i)_{1 \le i \le n}$$
$$= 0_{\mathbf{R}^n} + (x_i)_{1 \le i \le n}$$
$$= 0 + x.$$

And

$$x + 0 = (x_i)_{1 \le i \le n} + 0_{\mathbf{R}^n}$$
$$= (x_i + 0)_{1 \le i \le n}$$
$$= (x_i)_{1 \le i \le n}$$
$$= x.$$

Thus $x + 0 = 0 + x = x$.

(d) Additive inverse property.

$$x + (-x) = (x_i)_{1 \le i \le n} + (-1)(x_i)_{1 \le i \le n}$$
$$= (x_i)_{1 \le i \le n} + ((-1)x_i)_{1 \le i \le n}$$
$$= (x_i + (-1)x_i)_{1 \le i \le n}$$
$$= ((-1)x_i + x_i)_{1 \le i \le n}$$

$$= ((-1)x_i)_{1 \le i \le n} + (x_i)_{1 \le i \le n}$$
$$= (-1)(x_i)_{1 \le i \le n} + (x_i)_{1 \le i \le n}$$
$$= -x + x.$$

And

$$x + (-x) = (x_i)_{1 \le i \le n} + (-1)(x_i)_{1 \le i \le n}$$
$$= (x_i)_{1 \le i \le n} + ((-1)x_i)_{1 \le i \le n}$$
$$= (x_i + (-1)x_i)_{1 \le i \le n}$$
$$= (0)_{1 \le i \le n}$$
$$= 0.$$

Thus $x + (-x) = (-x) + x = 0$.

(e) Multiplicative associativity property.

$$(cd)x = ((cd)x_i)_{1 \le i \le n}$$
$$= (c(dx_i))_{1 \le i \le n}$$
$$= c(dx_i)_{1 \le i \le n}$$
$$= c(dx).$$

(f) Distributivity properties.

$$c(x + y) = c(x_i + y_i)_{1 \le i \le n}$$
$$= (c(x_i + y_i))_{1 \le i \le n}$$
$$= (cx_i + cy_i)_{1 \le i \le n}$$
$$= (cx_i)_{1 \le i \le n} + (cy_i)_{1 \le i \le n}$$
$$= c(x_i)_{1 \le i \le n} + c(y_i)_{1 \le i \le n}$$
$$= cx + cy.$$

And

$$(c + d)x = ((c + d)x_i)_{1 \le i \le n}$$
$$= (cx_i + dx_i)_{1 \le i \le n}$$
$$= (cx_i)_{1 \le i \le n} + (dx_i)_{1 \le i \le n}$$
$$= c(x_i)_{1 \le i \le n} + d(x_i)_{1 \le i \le n}$$

$$= cx + dx.$$

(g) Multiplicative identity property. Let $c = 1$ in Definition 6.1.1, then $1x = x$, as desired. $\qquad\square$

<div style="border:1px solid">

6.2

Derivatives in several variable calculus

</div>

Exercise 6.2.1. *Prove Lemma 6.2.1.*

> **Lemma 6.2.1.** Let E be a subset of $\mathbf{R}$, $f : E \to \mathbf{R}$ be a function, $x_0 \in E$, and $L \in \mathbf{R}$. Then the following two statements are equivalent.
>
> (a) f is differentiable at x_0, and $f'(x_0) = L$.
>
> (b) We have $\lim_{x \to x_0; x \in E \backslash \{x_0\}} \frac{f(x) - (f(x_0) + L(x - x_0))}{|x - x_0|} = 0$.

Proof. See Proposition 10.1.7 and Exercise 10.1.2, Analysis I. $\qquad\square$

Exercise 6.2.2. *Prove Lemma 6.2.4. (Hint: prove by contradiction. If $L_1 \neq L_2$, then there exists a vector v such that $L_1 v \neq L_2 v$; this vector must be non-zero (why?). Now apply the definition of derivative, and try to specialize to the case where $x = x_0 + tv$ for some scalar t, to obtain a contradiction.)*

> **Lemma 6.2.4** (Uniqueness of derivatives). Let E be a subset of $\mathbf{R}^n$, $f : E \to \mathbf{R}^m$ be a function, $x_0 \in E$ be an interior point of E, and let $L_1 : \mathbf{R}^n \to \mathbf{R}^m$ and $L_2 : \mathbf{R}^n \to \mathbf{R}^m$ be linear transformations. Suppose that f is differentiable at x_0 with derivative L_1, and also differentiable at x_0 with derivative L_2. Then $L_1 = L_2$.

Proof. Suppose for sake of contradiction that $L_1 \neq L_2$. there exists a non-zero v such that $L_1 v \neq L_2 v$. Since f is differentiable at x_0 with derivative

L_1 and L_2, by Definition 6.2.2, we have

$$\lim_{x \to x_0; x \in E \setminus \{x_0\}} \frac{\|f(x) - (f(x_0) + L_1(x - x_0))\|}{\|x - x_0\|} = 0,$$

and

$$\lim_{x \to x_0; x \in E \setminus \{x_0\}} \frac{\|f(x) - (f(x_0) + L_2(x - x_0))\|}{\|x - x_0\|} = 0.$$

Let $x = x_0 + tv$ for some $t \in \mathbf{R}_+$. Then we have

$$\lim_{t \to 0; t > 0, x_0 + tv \in E} \frac{\|f(x_0 + tv) - (f(x_0) + L_1(tv))\|}{\|tv\|} = 0,$$

and

$$\lim_{t \to 0; t > 0, x_0 + tv \in E} \frac{\|f(x_0 + tv) - (f(x_0) + L_2(tv))\|}{\|tv\|} = 0.$$

Then

$$\lim_{t \to 0; t > 0, x_0 + tv \in E} \frac{\|f(x_0 + tv) - f(x_0)\|}{\|tv\|} + \lim_{t \to 0; t > 0, x_0 + tv \in E} \frac{\|tL_1 v\|}{\|tv\|} = 0$$

and

$$\lim_{t \to 0; t > 0, x_0 + tv \in E} \frac{\|f(x_0 + tv) - f(x_0)\|}{\|tv\|} + \lim_{t \to 0; t > 0, x_0 + tv \in E} \frac{\|tL_2 v\|}{\|tv\|} = 0$$

Thus, we have

$$\lim_{t \to 0; t > 0, x_0 + tv \in E} \frac{\|tL_1 v\|}{\|tv\|} = \lim_{t \to 0; t > 0, x_0 + tv \in E} \frac{\|tL_2 v\|}{\|tv\|}.$$

This implies that $L_1 v = L_2 v$, a contradiction. $\qquad\square$

6.3

Partial and directional derivatives

Exercise 6.3.1. *Prove Lemma 6.3.5. (This will be similar to Exercise 6.2.1).*

> **Lemma 6.3.5.** Let E be a subset of $\mathbf{R}^n$, $f : E \to \mathbf{R}^m$ be a function, x_0 be an interior point of E, and let v be a vector in $\mathbf{R}^n$. If f is differentiable at x_0, then f is also differentiable in the direction v at x_0, and
>
> $$D_v f(x_0) = f'(x_0)v.$$

Proof. By Definition 6.2.2, we have

$$\lim_{x \to x_0 : x \neq x_0} \frac{\|f(x) - (f(x_0) + f'(x_0)(x - x_0))\|}{\|x - x_0\|} = 0.$$

Let $x = x_0 + tv$ for some $t \in \mathbf{R}$. Then

$$\lim_{t \to 0 : t > 0} \frac{\|f(x_0 + tv) - (f(x_0) + f'(x_0)(tv))\|}{\|tv\|} = 0.$$

Because $\|v\| > 0$, above limit implies that

$$\lim_{t \to 0 : t > 0} \frac{f(x_0 + tv) - f(x_0)}{t} = f'(x_0)v.$$

Thus $D_v f(x_0) = f'(x_0)v$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Exercise 6.3.2. *Let E be a subset of $\mathbf{R}^n$, let $f : E \to \mathbf{R}^m$ be a function, let x_0 be an interior point of E, and let $1 \leq j \leq n$. Show that $\frac{\partial f}{\partial x_j}(x_0)$ exists if and only if $D_{e_j} f(x_0)$ and $D_{-e_j} f(x_0)$ exist and are negatives of each other (thus $D_{e_j} f(x_0) = -D_{-e_j} f(x_0)$); furthermore, one has $\frac{\partial f}{\partial x_j}(x_0) = D_{e_j} f(x_0)$ in this case.*[1]

Proof. Suppose that $\frac{\partial f}{\partial x_j}(x_0)$ exists, by Definition 6.3.7, we have

$$\frac{\partial f}{\partial x_j}(x_0) = \lim_{t \to 0; t \neq 0, x_0 + te_j \in E} \frac{f(x_0 + te_j) - f(x_0)}{t}.$$

By Definition 6.3.1, when $t > 0$, we have

$$D_{e_j} f(x_0) = \lim_{t \to 0; t > 0} \frac{f(x_0 + te_j) - f(x_0)}{t} = \frac{\partial f}{\partial x_j}(x_0),$$

when $t < 0$, we have

$$D_{-e_j} f(x_0) = -\lim_{t \to 0; -t > 0} \frac{f(x_0 + te_j) - f(x_0)}{t} = -\frac{\partial f}{\partial x_j}(x_0).$$

Thus $D_{e_j} f(x_0) = -D_{-e_j} f(x_0)$.

Conversely, suppose that $D_{e_j} f(x_0)$ and $D_{-e_j} f(x_0)$ exist and $D_{e_j} f(x_0) = -D_{-e_j} f(x_0)$. By Definition 6.3.1, we have

$$D_{e_j} f(x_0) = \lim_{t \to 0; t > 0} \frac{f(x_0 + te_j) - f(x_0)}{t}$$

---

[1] In Exercise 6.3.2, $D_{e_j} f(x_0) = D_{-e_j} f(x_0)$ should be $D_{e_j} f(x_0) = -D_{-e_j} f(x_0)$. — Errata from Tao.

and
$$D_{-e_j} f(x_0) = \lim_{t \to 0; t > 0} \frac{f(x_0 - te_j) - f(x_0)}{t}.$$
By Lemma 6.3.6, we have $D_{e_j} f(x_0) = f'(x_0)e_j$ and $D_{-e_j} f(x_0) = -f'(x_0)e_j$. Thus, $D_{e_j} f(x_0) = -D_{-e_j} = f'(x_0)e_j$.

We need to show that $\frac{f(x_0 + te_j) - f(x_0)}{t}$ converges to $f'(x_0)e_j$. When $t > 0$, we have
$$\lim_{t \to 0; t > 0, x_0 + te_j \in E} \frac{f(x_0 + te_j) - f(x_0)}{t} = f'(x_0)e_j.$$
When $t < 0$,
$$\lim_{t \to 0; -t > 0, x_0 - te_j \in E} \frac{f(x_0 - te_j) - f(x_0)}{-t} = -f'(x_0)e_j,$$
so that
$$\lim_{t \to 0; -t > 0, x_0 - te_j \in E} \frac{f(x_0 - te_j) - f(x_0)}{t} = f'(x_0)e_j.$$
Since left limit equals to right limit,
$$\lim_{t \to 0; t \neq 0} \frac{f(x_0 + te_j) - f(x_0)}{t}$$
exists and equals to $f'(x_0)e_j$. Thus $\frac{\partial f}{\partial x_j}(x_0) = D_{e_j} f(x_0)$. $\qquad \square$

Exercise 6.3.3. *Let $f : \mathbf{R}^2 \to \mathbf{R}$ be the function defined by $f(x, y) := \frac{x^3}{x^2 + y^2}$ when $(x, y) \neq (0, 0)$, and $f(0, 0) := 0$. Show that f is not differentiable at $(0, 0)$, despite being differentiable in every direction $v \in \mathbf{R}^2$ at $(0, 0)$. Explain why this does not contradict Theorem 6.3.8.*

Proof. Let $v = (v_1, v_2)$ such that $(x, y) = (tv_1, tv_2)$. We need to verify Definition 6.3.1:
$$\lim_{t \to 0; t > 0} \frac{f(x, y) - f(0, 0)}{t} = \lim_{t \to 0; t > 0} \frac{1}{t} \cdot \frac{(tv_1)^3}{(tv_1)^2 + (tv_2)^2} = \frac{v_1^3}{v_1^2 + v_2^2}.$$
Thus f is differentiable in every direction $v \in \mathbf{R}^2$ at $(0, 0)$.

Taking v to be e_1 and e_2, we have
$$\frac{\partial f}{\partial x}(0, 0) = 1, \qquad \frac{\partial f}{\partial y}(0, 0) = 0.$$
For $(x, y) \neq (0, 0)$, we have
$$\frac{\partial f}{\partial x}(x, y) = \frac{x^2(x^2 + 3y^2)}{(x^2 + y^2)^2}, \qquad \frac{\partial f}{\partial y}(x, y) = \frac{-2yx^3}{(x^2 + y^2)^2}.$$

Thus, we can verify that $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ is continuous at $(0,0)$. By Lemma 6.3.5 and Theorem 6.3.8, we have

$$D_{(x,y)}f(0,0) = x\frac{\partial f}{\partial x}(0,0) + y\frac{\partial f}{\partial y}(0,0) = x.$$

Now turn to show that f is not differentiable at $(0,0)$. Suppose for sake of contradiction that f is differentiable at $(0,0)$ with derivative L, where $L : \mathbf{R}^2 \to \mathbf{R}$ is linear transformation. By Lemma 6.3.5, we have $L(x,y) = D_{(x,y)}f(0,0) = x$. We have to verify Definition 6.2.2 that

$$\lim_{(x,y)\to(0,0):(x,y)\neq(0,0)} \frac{\left\|\frac{x^3}{x^2+y^2} - L(x,y)\right\|}{\|(x,y)\|} = 0.$$

Since

$$\frac{\left\|\frac{x^3}{x^2+y^2} - L(x,y)\right\|}{\|(x,y)\|} = \frac{\left\|\frac{x^3}{x^2+y^2} - x\right\|}{\sqrt{x^2+y^2}} = \frac{y^2}{(x^2+y^2)^{3/2}}$$

If we let $x = y$, then $\lim_{(x,x)\to(0,0):(x,x)\neq(0,0)} \frac{xy^2}{(x^2+y^2)^{3/2}} = 2^{-3/2}$, a contradiction. Thus f is not differentiable at $(0,0)$. $\qquad\square$

Exercise 6.3.4. *Let $f : \mathbf{R}^n \to \mathbf{R}^m$ be a differentiable function such that $f'(x) = 0$ for all $x \in \mathbf{R}^n$. Show that f is constant. (Hint: you may use the mean-value theorem or fundamental theorem of calculus for one-dimensional functions, but bear in mind that there is no direct analogue of these theorems for several-variable functions. I would not advise proceeding via first principles.) For a tougher challenge, replace the domain $\mathbf{R}^n$ by an open connected subset Ω of $\mathbf{R}^n$.*

CHAPTER VII

Lebesgue measure

The goal: Lebesgue measure

(i) (Borel property) Every open set in $\mathbf{R}^n$ is measurable, as is every closed set.

(ii) (Complementarity) If Ω is measurable, then $\mathbf{R}^n \setminus \Omega$ is also measurable.

(iii) (Boolean algebra property) If $(\Omega_j)_{j \in J}$ is any finite collection of measurable sets (so J is finite), then the union $\bigcup_{j \in J} \Omega_j$ and intersection $\bigcap_{j \in J} \Omega_j$ are also measurable.

(iv) (σ-algebra property) If $(\Omega_j)_{j \in J}$ are any countable collection of measurable sets (so J is countable), then the union $\bigcup_{j \in J} \Omega_j$ and intersection $\bigcap_{j \in J} \Omega_j$ are also measurable.

(v) (Empty set) The empty set $\emptyset$ has measure $m(\emptyset) = 0$.

(vi) (Positivity) We have $0 \leq m(\Omega) \leq +\infty$ for every measurable set Ω.

(vii) (Monotonicity) If $A \subseteq B$, and A and B are both measurable, then $m(A) \leq m(B)$.

(viii) (Finite sub-additivity) If $(A_j)_{j \in J}$ are a finite collection of measurable sets, then $m(\bigcup_{j \in J} A_j) \leq \sum_{j \in J} m(A_j)$.

(ix) (Finite additivity) If $(A_j)_{j \in J}$ are a finite collection of *disjoint* measurable sets, then $m(\bigcup_{j \in J} A_j) = \sum_{j \in J} m(A_j)$.

(x) (Countable sub-additivity) If $(A_j)_{j \in J}$ are a countable collection of measurable sets, then $m(\bigcup_{j \in J} A_j) \leq \sum_{j \in J} m(A_j)$.

(xi) (Countable additivity) If $(A_j)_{j \in J}$ are a countable collection of *disjoint* measurable sets, then $m(\bigcup_{j \in J} A_j) = \sum_{j \in J} m(A_j)$.

(xii) (Normalization) The unit cube $[0,1]^n = \{(x_1, \cdots, x_n) \in \mathbf{R}^n : 0 \leq x_j \leq 1$ for all $1 \leq j \leq n\}$ has measure $m([0,1]^n) = 1$.

(xiii) (Translation invariance) If Ω is a measurable set, and $x \in \mathbf{R}^n$, then $x + \Omega := \{x + y : y \in \Omega\}$ is also measurable, and $m(x + \Omega) = m(\Omega)$.

7.2 First attempt: Outer measure

Exercise 7.2.1. *Prove Lemma 7.2.5. (Hint: you will have to use the definition of* inf*, and probably introduce a parameter* ε*. You may have to treat separately the cases when certain outer measures are equal to* $+\infty$*. (viii) can be deduced from (x) and (v). For (x), label the index set* J *as* $J = \{j_1, j_2, j_3, \cdots\}$*, and for each* A_j*, pick a covering of* A_j *by boxes whose total volume is no larger than* $m^*(A_j) + \varepsilon/2^j$*.)*

Lemma 7.2.5 (Properties of outer measure). Outer measure has the following six properties:

(v) (Empty set) The empty set $\emptyset$ has outer measure $m^*(\emptyset) = 0$.

(vi) (Positivity) We have $0 \leq m^*(\Omega) \leq +\infty$ for every measurable set Ω.

(vii) (Monotonicity) If $A \subseteq B \subseteq \mathbf{R}^n$, then $m^*(A) \leq m^*(B)$.

(viii) (Finite sub-additivity) If $(A_j)_{j \in J}$ are a finite collection of subsets of $\mathbf{R}^n$, then $m^*(\bigcup_{j \in J} A_j) \leq \sum_{j \in J} m^*(A_j)$.

(x) (Countable sub-additivity) If $(A_j)_{j \in J}$ are a countable collection of subsets of $\mathbf{R}^n$, then $m^*(\bigcup_{j \in J} A_j) \leq \sum_{j \in J} m^*(A_j)$.

(xiii) (Translation invariance) If Ω is a subset of $\mathbf{R}^n$, and $x \in \mathbf{R}^n$, then $m^*(x + \Omega) = m^*(\Omega)$.

Proof. (v). Since empty set is subset of any set, in particular, we have $B_j = \emptyset$ for all $j \in J$. Thus by Definition 7.2.1, 7.2.4, $m^*(\emptyset) = 0$ for that $\sum_{j \in J} \text{vol}(\emptyset)$ is non-negative.

(vi). Since $\text{vol}(B_j)$ is non-negative for all $j \in J$, we have $0 \leq m^*(\Omega)$. If B_j is bounded, there exists a real number $r \in \mathbf{R}$ such that $\sum_{j \in J} \text{vol}(B_j) = r$. While if B_j is unbounded for all $j \in J$, then for every r, we have $r < \sum_{j \in J} \text{vol}(B_j)$, thus $m^*(\Omega) = +\infty$.

(vii). If $A \subseteq B$, then for every collection of boxes cover B must cover A, thus $m^*(A) \leq m^*(B)$.

(viii). This is a special case of (x).

(x). If $m^*(A_j) = +\infty$ for some $j \in J$, the conclusion is trivial. Here, we suppose that $m^*(A_j)$ is finite for all $j \in J$.[1] Let $\varepsilon > 0$, for each $j \in J$ there exists a cover such that
$$A_j \subseteq \bigcup_{n \in \mathbf{N}} A_{j_n}.$$
Then for each A_j, there is a j such that
$$\sum_{n \in \mathbf{N}} A_{j_n} < m^*(A_j) + \frac{\varepsilon}{2^j}.$$
Thus by Definition 7.2.4,
$$m^*\left(\bigcup_{j \in J} A_j\right) \leq \sum_{j \in J} \sum_{n \in \mathbf{N}} A_{j_n}$$
$$< \sum_{j \in J} \left(m^*(A_j) + \frac{\varepsilon}{2^j}\right)$$

---

[1] Recall Definition 6.2.1, *Analysis I*, that x is finite iff it is a real number, and infinite iff it is equal to $+\infty$ or $-\infty$.

$$= \sum_{j \in J} m^*(A_j) + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, as ε converges to zero, we have

$$m^*\left(\bigcup_{j \in J} A_j\right) \leq \sum_{j \in J} m^*(A_j).$$

(xiii). Suppose that $m^*(A_j)$ is finite for all $j \in J$. There exists a cover such that $\Omega \subseteq \bigcup_{j \in J} B_j$. Then for $x \in \mathbf{R}^n$, we have $\Omega + x \subseteq \bigcup_{j \in J}(B_j + x)$. Since for each $y \in B_j + x$, so that $y_i \in (a_i + x_i, b_i + x_i)$, this implies that $\mathrm{vol}(B_j + x) = \mathrm{vol}(B_j)$. Therefore, $m^*(x + \Omega) = m^*(\Omega)$. □

Exercise 7.2.2. *Let A be a subset of $\mathbf{R}^n$, and let B be a subset of $\mathbf{R}^m$. Note that the Cartesian product $\{(a,b) : a \in A, b \in B\}$ is then a subset of $\mathbf{R}^{n+m}$. Show that $m^*_{n+m}(A \times B) \leq m^*_n(A) m^*_m(B)$. (It is in fact true that $m^*_{n+m}(A \times B) = m^*_n(A) m^*_m(B)$, but this is substantially harder to prove.) Here we adopt the convention that $c \times +\infty = +\infty \times c$ is infinite for any $0 < c \leq +\infty$ and vanishes for $c = 0$.*

In Exercises 7.2.3-7.2.5, we assume that $\mathbf{R}^n$ is a Euclidean space, and we have a notion of measurable set in $\mathbf{R}^n$ (which may or may not coincide with the notion of Lebesgue measurable set) and a notion of measure (which may or may not co-incide with Lebesgue measure) which obeys axioms (i)-(xiii).

Exercise 7.2.3.

(a) *Show that if $A_1 \subseteq A_2 \subseteq A_3 \cdots$ is an increasing sequence of measurable sets (so $A_j \subseteq A_{j+1}$ for every positive integer j), then we have $m(\bigcup_{j=1}^{\infty} A_j) = \lim_{j \to \infty} m(A_j)$.*

(b) *Show that if $A_1 \supseteq A_2 \supseteq A_3 \cdots$ is a decreasing sequence of measurable sets (so $A_j \supseteq A_{j+1}$ for every positive integer j), and $m(A_1) < +\infty$, then we have $m(\bigcap_{j=1}^{\infty} A_j) = \lim_{j \to \infty} m(A_j)$.*

Proof. (a) Let $B_j = A_j \setminus A_{j-1}$, then for every positive integer j, B_j are disjoint. Since $A_j = \bigcup_{n=1}^{j} B_n$ and $\bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^{\infty} B_j$, by (ix) and (xi), we have

$$m\left(\bigcup_{j=1}^{\infty} A_j\right) = m\left(\bigcup_{j=1}^{\infty} B_j\right)$$

$$= \sum_{j=1}^{\infty} m(B_j)$$

$$= \lim_{j\to\infty} \sum_{n=1}^{j} m(B_n)$$

$$= \lim_{j\to\infty} m\left(\bigcup_{n=1}^{j} B_n\right)$$

$$= \lim_{j\to\infty} m(A_j).$$

(b) Let $B_j = A_1 \setminus A_j$. Then we have $B_j \subseteq B_{j+1}$ for every positive integer j. By (a), we have

$$m\left(\bigcup_{j=1}^{\infty} B_j\right) = m\left(\bigcup_{j=1}^{\infty}(A_1 \setminus A_j)\right) = \lim_{j\to\infty} m(A_1 \setminus A_j).$$

Since $\bigcup_{j=1}^{\infty}(A_1 \setminus A_j) = A_1 \setminus \bigcap_{j=1}^{\infty} A_j$, we have

$$m\left(\bigcup_{j=1}^{\infty} B_j\right) = m\left(A_1 \setminus \bigcap_{j=1}^{\infty} A_j\right).$$

This is easy to see that $\bigcap_{j=1}^{\infty} A_j \subseteq A_1$ and $A_j \subseteq A_1$. Since $A_1 \setminus A_j$ and A_j are disjoint, so is $A_1 \setminus \bigcap_{j=1}^{\infty} A_j$ and A_j, from (xi), we have

$$m(A_1 \setminus A_j) = m(A_1) - m(A_j)$$

and

$$m\left(A_1 \setminus \bigcap_{j=1}^{\infty} A_j\right) = m(A_1) - m\left(\bigcap_{j=1}^{\infty} A_j\right).$$

Then

$$\lim_{j\to\infty}\left(m(A_1) - m(A_j)\right) = m(A_j) - \lim_{j\to\infty} m(A_j) = m(A_1) - m\left(\bigcap_{j=1}^{\infty} A_j\right).$$

Thus, $m(\bigcap_{j=1}^{\infty} A_j) = \lim_{j\to\infty} m(A_j)$. $\qquad\square$

Remark 7.2.3. In probability theory, $A_j \subseteq A_{j+1}$ and $A_j \supseteq A_{j+1}$ for every positive integer j is usually denoted by $A_j \uparrow$ and $A_j \downarrow$. Then assertions (a) and (b) are represented as: if $A_j \uparrow A$ or $A_j \downarrow A$, then $m(A_j) \to m(A)$, where $A = \bigcup_{j=1}^{\infty} A_j$ in (a) and $A = \bigcap_{j=1}^{\infty} A_j$ in (b).

7.3

Outer measure is not additive

THERE IS NO EXERCISES IN SECTION 7.3.

7.4

Measurable sets

Exercise 7.4.1. *If A is an open interval in $\mathbf{R}$, show that $m^*(A) = m^*(A \cap (0, \infty)) + m^*(A \setminus (0, \infty))$.*

Proof. Let A takes the form of (a, b) where $b > a$ and $a, b \in \mathbf{R}$. Then $m^*(A) = b - a$. If $A \cap (0, \infty) = \emptyset$, then

$$m^*(A \cap (0, \infty)) + m^*(A \setminus (0, \infty)) = m^*(\emptyset) + m^*(A) = m(A).$$

If $A \cap (0, \infty) = (a, b)$, then

$$m^*(A \cap (0, \infty)) + m^*(A \setminus (0, \infty)) = m^*(A) + m^*(\emptyset) = m(A).$$

While if $A \cap (0, \infty) \subsetneq (a, b)$, then

$$
\begin{aligned}
m^*(A \cap (0, \infty)) + m^*(A \setminus (0, \infty)) &= m^*((0, b)) + m^*((a, 0]) \\
&= (b - 0) + (0 - a) \\
&= m^*(A).
\end{aligned}
$$

Thus $m^*(A) = m^*(A \cap (0, \infty)) + m^*(A \setminus (0, \infty))$ is hold for arbitrary open interval $A \subseteq \mathbf{R}$. $\qquad\square$

Exercise 7.4.2. *If A is an open box in $\mathbf{R}^n$, and E is the half-plane $E := \{(x_1, \cdots, x_n) \in \mathbf{R}^n : x_n > 0\}$, show that $m^*(A) = m^*(A \cap E) + m^*(A \setminus E)$. (Hint: use Exercise 7.4.1.)*

Proof. Let B_i to be an open box in $\mathbf{R}$, and $F_i := \{x \in \mathbf{R} : x > 0\}$ for all positive integer i. This is easy to see that $A_n = \prod_{i=1}^{n} B_i$ and $E_n = \prod_{i=1}^{n} F_i$ where A_n is an open box and E_n is the half-plane in $\mathbf{R}^n$. We use induction on n to prove that

$$m^*(A) \geq m^*(A \cap E) + m^*(A \setminus E)$$

When $n = 1$, inequality is hold for A_1 by Exercise 7.4.1. Now we suppose inductively that $m^*(A_{n-1}) \geq m^*(A_{n-1} \cap E_{n-1}) + m^*(A_{n-1} \setminus E_{n-1})$. Then

$$
\begin{aligned}
m_n^*(A) =& m_n^*(A_{n-1} \times B_n) \\
=& m_{n-1}^*(A_{n-1}) m_1^*(B_n) \\
\geq& \left(m_{n-1}^*(A_{n-1} \cap E_{n-1}) + m_{n-1}^*(A_{n-1} \setminus E_{n-1}) \right) \\
& \times \left(m_1^*(B_n \cap F_n) + m_1^*(B_n \setminus F_n) \right) \\
\geq& m_n^*((A_{n-1} \times B_n) \cap (E_{n-1} \times F_n)) \\
& + m_n^*((A_{n-1} \times B_n) \setminus (E_{n-1} \times F_n)) \\
=& m_n^*(A \cap E) + m_n^*(A \setminus E).
\end{aligned}
$$

This close the induction and we complete the half of the proof. By (x), we always have $m^*(A) \leq m^*(A \cap E) + m^*(A \setminus E)$. Thus $m^*(A) = m^*(A \cap E) + m^*(A \setminus E)$. $\qquad\square$

Exercise 7.4.3. *Prove Lemma 7.4.2. (Hint: use Exercise 7.4.2.)*

> **Lemma 7.4.2** (Half-spaces are measurable). The half-space
> $$\{(x_1, \cdots, x_n) \in \mathbf{R}^n : x_n > 0\}$$
> is measurable.

Proof. Denote the half-space by E and let A be any subset of $\mathbf{R}^n$. Given a cover $(B_j)_{j \in J}$ of A for at most countable J. Then there is an open cover $\bigcup_{j \in J} B_j'$ such that $\bigcup_{j \in J} B_j \subseteq B := \bigcup_{j \in j} B_j'$. By Exercise 7.4.2, we have

$$m^*(B) = m^*(B \cap E) + m^*(B \setminus E) \geq m^*(A \cap E) + m^*(A \setminus E).$$

Since $m^*(A)$ is the infimum of $\sum_{j \in J} \operatorname{vol}(B_j')$, we have $m^*(A) \geq m^*(A \cap E) + m^*(A \setminus E)$. Other hand of proof is immediately from (x). $\qquad\square$

Exercise 7.4.4. *Prove Lemma 7.4.2. (Hint: for (c), first prove that*

$$m^*(A) = m^*(A \cap E_1 \cap E_2) + m^*(A \cap E_1 \setminus E_2)$$
$$+ m^*(A \cap E_2 \setminus E_1) + m^*(A \setminus (E_1 \cup E_2)).$$

A Venn diagram may be helpful. Also you may need the finite sub-additivity property. Use (c) to prove (d), and use (bd) and the various versions of Lemma 7.4.2 to prove (e)).

> **Lemma 7.4.4** (Properties of measurable sets).
>
> (a) If E is measurable, then $\mathbf{R}^n \setminus E$ is also measurable.
>
> (b) (Translation invariance) If E is measurable, and $x \in \mathbf{R}^n$, then $x + E$ is also measurable, and $m(x + E) = m(E)$.
>
> (c) If E_1 and E_2 are measurable, then $E_1 \cap E_2$ and $E_1 \cup E_2$ are measurable.
>
> (d) (Boolean algebra property) If $E_1, E_2, \cdots, E_N$ are measurable, then $\bigcup_{j=1}^{N} E_j$ and $\bigcap_{j=1}^{N} E_j$ are measurable.
>
> (e) Every open box, and every closed box, is measurable.
>
> (f) Any set E of outer measure zero (i.e., $m^*(E) = 0$) is measurable.

Proof. (a) Let A is arbitrary set. Since E is measurable, we have

$$m^*(A) = m^*(A \cap E) + m^*(A \setminus E).$$

Since $A \cap (\mathbf{R}^n \setminus E) = A \setminus$ and $A \setminus (\mathbf{R}^n \setminus E) = A \cap E$, we have

$$m^*(A) = m^*(A \cap (\mathbf{R}^n \setminus E)) + m^*(A \setminus (\mathbf{R}^n \setminus E)).$$

Thus $(\mathbf{R}^n \setminus E)$ is measurable.

(b) This immediately comes from Lemma 7.2.5(xiii).

(c) Since E_1 and E_2 is measurable, for arbitrary set A, we have $A \cap E_1$ and $A \setminus E_1$ satisfying that

$$m^*(A \cap E_1) = m^*(A \cap E_1 \cap E_2) + m^*(A \cap E_1 \setminus E_2)$$

and

$$m^*(A \setminus E_1) = m^*(A \setminus E_1 \cap E_2) + m^*(A \setminus (E_1 \cup E_2)).$$

Since
$$m^*(A) = m^*(A \cap E_1) + m^*(A \setminus E_1),$$
we obtain that
$$m^*(A) = m^*(A \cap E_1 \cap E_2) + m^*(A \cap E_1 \setminus E_2)$$
$$+ m^*(A \cap E_2 \setminus E_1) + m^*(A \setminus (E_1 \cup E_2)).$$
Here, we replace A with $A \cap (E_1 \cup E_2)$, then
$$m^*(A \cap (E_1 \cup E_2)) = m^*(A \cap E_1 \cap E_2) + m^*(A \cap E_1 \setminus E_2)$$
$$+ m^*(A \cap E_2 \setminus E_1) + 0.$$
Notice that we have following relations:
$$A \cap (E_1 \cup E_2) \cap E_1 \cap E_2 = A \cap E_1 \cap E_2,$$
$$A \cap (E_1 \cup E_2) \cap E_1 \setminus E_2 = A \cap E_1 \setminus E_2,$$
$$A \cap (E_1 \cup E_2) \cap E_2 \setminus E_1 = A \cap E_2 \setminus E_1,$$
$$A \cap (E_1 \cup E_2) \setminus (E_1 \cup E_2) = \emptyset.$$
Thus we have
$$m^*(A) = m^*(A \cap (E_1 \cup E_2)) + m^*(A \setminus (E_1 \cup E_2)),$$
and $E_1 \cup E_2$ is measurable.

Again, replace A with $A \setminus (E_1 \cap E_2)$. Then we have
$$A \setminus (E_1 \cap E_2) \cap E_1 \cap E_2 = \emptyset,$$
$$A \setminus (E_1 \cap E_2) \cap E_1 \setminus E_2 = A \cap E_1 \setminus E_2,$$
$$A \setminus (E_1 \cap E_2) \cap E_2 \setminus E_1 = A \cap E_2 \setminus E_1,$$
$$A \setminus (E_1 \cap E_2) \setminus (E_1 \cup E_2) = A \setminus (E_1 \cup E_2).$$

Hence
$$m^*(A) = m^*(A \cap (E_1 \cap E_2)) + m^*(A \setminus (E_1 \cap E_2)),$$
and $E_1 \cap E_2$ is measurable.

(d) Use induction and conclusions follow immediately.

(e) Let
$$E := \{(x_1, \cdots, x_n) \in \mathbf{R}^n : x_n > 0\}$$

and

$$F := \{(x_1, \cdots, x_n) \in \mathbf{R}^n : x_n < 0\}.$$

By Lemma 7.4.2, E and F are measurable. Then for open box B and $a_i, b_i \in \mathbf{R}$ such that $a_i \leq b_i$ for all $1 \leq i \leq n$, we have

$$B = \bigcap_{i=1}^{n} ((E + b_i) \cap (F + a_i))$$

By Lemma 7.4.4(bd), B is measurable. This is similar to show that closed box is measurable.

(f) Let A be an arbitrary set, we need to show that

$$m^*(A) = m^*(A \cap E) + (A \setminus E).$$

Obviously, we have $m^*(A) \leq m^*(A \cap E) + (A \setminus E)$. Consider the other hand. This is easy to see that $m^*(E) \geq m^*(A \cap E)$, since $m^*(E) = 0$, we have $m^*(A \cap E) = 0$. Hence, $m^*(A) \geq m^*(A \setminus E)$, as desired.   □

Exercise 7.4.5. *Show that the set E used in the proof of Propositions 7.3.1 and 7.3.3 is non-measurable.*

Exercise 7.4.6. *Prove Lemma 7.4.5.*

> **Lemma 7.4.5** (Finite additivity). If $(E_j)_{j \in J}$ are a finite collection of disjoint measurable sets and any set A (not necessarily measurable), we have
>
> $$m^* \left(A \cap \bigcup_{j \in J} E_j \right) = \sum_{j \in J} m^*(A \cap E_j).$$
>
> Furthermore, we have $m(\bigcup_{j \in J} E_j) = \sum_{j \in J} m(E_j)$.

Proof. Suppose that J has cardinality with n. We use induction on n. Base case is trivial. Now suppose inductively that the equation is hold for $n - 1$, we need to show the case of n. Because $(E_j)_{j \in J}$ are disjoint, we have

$$A \cap \left(\bigcup_{j=1}^{n} E_j \right) \cap E_n = A \cap E_n$$

and

$$A \cap \left(\bigcup_{j=1}^{n} E_j \right) \setminus E_n = A \cap \bigcup_{j=1}^{n-1} E_j.$$

Then, by inductive hypothesis, we have

$$m^* \left(A \cap \bigcup_{j=1}^{n} E_j \right) = m^* \left(A \cap \bigcup_{j=1}^{n} E_j \cap E_n \right) + m^* \left(A \cap \bigcup_{j=1}^{n} E_j \setminus E_n \right)$$

$$= m^*(A \cap E_n) + m^* \left(A \cap \bigcup_{j=1}^{n-1} E_j \right)$$

$$= m^*(A \cap E_n) + \sum_{j=1}^{n-1} m^*(A \cap E_j)$$

$$= \sum_{j=1}^{n} m^*(A \cap E_j).$$

Since A is an arbitrary set, let $A = \bigcup_{j=1}^{n} E_j$, we have $m(\bigcup_{j \in J} E_j) = \sum_{j \in J} m(E_j)$. $\qquad \square$

Exercise 7.4.7. *Use Lemma 7.4.5 to prove Corollary 7.4.7.*

> **Corollary 7.4.7.** If $A \subseteq B$ are two measurable sets, then $B \setminus A$ is also measurable, and
> $$m(B \setminus A) = m(B) - m(A).$$

Proof. Since A and B are measurable, and $B \setminus A = B \cap (\mathbf{R}^n \setminus A)$, then $B \setminus A$ is measurable by Lemma 7.4.4(ac). Since $B \setminus A$ and A are disjoint, by Lemma 7.4.5, we have $m(A) + m(B \setminus A) = m(B)$. Thus $m(B \setminus A) = m(B) - m(A)$. $\qquad \square$

Exercise 7.4.8. *Prove Lemma 7.4.9. (Hint: for the countable union problem, write $J = \{j_1, j_2, \cdots\}$, write $F_N := \bigcup_{k=1}^{N} \Omega_{j_k}$, and write $E_N := F_N \setminus F_{N-1}$, with the understanding that F_0 is the empty set. Then apply Lemma 7.4.8. For the countable intersection problem, use what you just did and Lemma 7.4.4(a).)*

Lemma 7.4.9 (σ-algebra property). If $(\Omega_j)_{j \in J}$ are any countable collection of measurable sets (so J is countable), then the union $\bigcup_{j \in j} \Omega_j$ and the intersection $\bigcap_{j \in j} \Omega_j$ are also measurable.

Proof. Let $\Omega := \bigcup_{j \in J} \Omega_j$, let A be an arbitrary set. We need to show that
$$m^*(A) = m^*(A \cap \Omega) + m^*(A \setminus \Omega).$$

Since J is countable, we write $J = \{j_1, j_2, \cdots\}$. Let F_N be the set $F_N := \bigcup_{k=1}^{N} \Omega_{j_k}$, and let $E_N := F_N \setminus F_{N-1}$. Then E_N are disjoint for all $N \geq 1$, by Lemma 7.4.8, we have
$$\bigcup_{N=1}^{\infty} E_N = \bigcup_{N=1}^{\infty} (F_N \setminus F_{N-1})$$
$$= \bigcup_{j \in J} \Omega_j$$

is measurable. Thus $\bigcup_{j \in J} \Omega_j$ is measurable.

By Lemma 7.4.4(a), we have $\bigcup_{j \in J} \Omega_j = \mathbf{R}^n \setminus \bigcup_{j \in J} \Omega_j$ is measurable. $\square$

7.5

Measurable functions

Exercise 7.5.1. *Prove Lemma 7.5.3. (Hint: use Lemma 7.4.10 and the σ-algebra property.)*

Lemma 7.5.3. Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f : \Omega \to \mathbf{R}^m$ be a function. Then f is measurable if and only if $f^{-1}(B)$ is measurable for every open box B.

Proof. Suppose that f is measurable. By Definition 7.5.1, $f^{-1}(V)$ is measurable for every open set $V \subseteq \mathbf{R}^m$. This implies that $f^{-1}(B)$ is measurable for every open box B.

Conversely, suppose that $f^{-1}(B)$ is measurable for every open box B. We write B as a sequence of open boxes $(B_j)_{j \in J}$, where J is finite or countable. $(B_j)_{j \in J}$ is measurable by Lemma 7.4.11. By Lemma 7.4.10, every open set $V \subseteq \mathbf{R}^n$ can be represented as $V = \bigcup_{j \in J} B_j$. Since $f^{-1}(B_j)$

is measurable for $(B_j)_{j \in J}$, we have $f^{-1}(V) = f^{-1}(\bigcup_{j \in J} B_j) = \bigcup_{j \in J} f^{-1}(B_j)$ for every $V \subseteq \mathbf{R}^m$, as desired. $\square$

Exercise 7.5.2. *Use Lemma 7.5.3 to deduce Corollary 7.5.4.*

> **Corollary 7.5.4.** Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f : \Omega \to \mathbf{R}^m$ be a function. Suppose that $f = (f_1, \cdots, f_m)$, where $f_j : \Omega \to \mathbf{R}$ is the jth co-ordinate of f. Then f is measurable if and only if all of the f_j are individually measurable.

Bibliography

[1] T. Tao, *Analysis*, Vol. II, Third Edition, Hindustan Book Agency, 2015.

[2] W. Rudin, *Principles of Mathematical Analysis*, Third Edition, MacGraw Hill, 1976.

[3] V.A. Zorich, *Mathematical Analysis*, Vol. II, Second Edition, Springer-Verlag Berlin Heidelberg, 2016.

[4] H.L. Royden and P.M. Fitzpatrick, *Real Analysis*, Fourth Edition, China Machine Press, 2010.

[5] J. Munkres, *Topology*, Second Edition, Pearson, 2014.