A

Metrics, Norms, Inner Products, and Topology

These appendices collect the background material needed for the main part of the volume. In keeping with the philosophy of this text, we formulate these as "mini-courses" with the goal of providing substantial, though not exhaustive, introductions to and reviews of their respective subjects. Topics that are not typically part of standard beginning mathematics graduate courses are given more detailed attention, while other results are either formulated as exercises (often with hints) or stated without proof. Sources for additional information on the material of this and most of the other appendices include Folland's real analysis text [Fol99], Conway's functional analysis text [Con90], and the the operator theory/Hilbert space text [GG01] by Gohberg and Goldberg.

A.1 Notational Conventions

We first review some of the notational conventions that are used throughout this volume.

Unless otherwise specified, all vector spaces are taken over the complex field $\mathbb{C}$. In particular, functions whose domain is $\mathbb{R}^d$ (or a subset of $\mathbb{R}^d$) are generally allowed to take values in the complex plane $\mathbb{C}$.

Integrals with unspecified limits are taken over either the real line or $\mathbb{R}^d$, according to context. In particular, if $f\colon \mathbb{R} \to \mathbb{C}$, then we take

$$\int f(x)\,dx \;=\; \int_{-\infty}^{\infty} f(x)\,dx.$$

The extended real line is $\mathbb{R} \cup \{-\infty, \infty\} = [-\infty, \infty]$. We use the conventions that $1/0 = \infty$, $1/\infty = 0$, and $0 \cdot \infty = 0$.

If $1 \le p \le \infty$ is given, then its *dual index* or *dual exponent* is the extended real number p' that satisfies

$$\frac{1}{p} + \frac{1}{p'} \;=\; 1.$$

Explicitly,

$$p' = \frac{p}{p-1}.$$

The dual index lies in the range $1 \le p' \le \infty$, and we have $1' = \infty$, $2' = 2$, and $\infty' = 1$.

The *Kronecker delta* is

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \ne j. \end{cases}$$

A.2 Metrics and Convergence

A metric determines a notion of distance between points in a set.

Definition A.1 (Metric Space). Let X be a set. A *metric* on X is a function $\mathrm{d}\colon X \times X \to \mathbb{R}$ such that for all $f,\ g,\ h \in X$ we have:

(a) $\mathrm{d}(f,g) \ge 0$,

(b) $\mathrm{d}(f,g) = 0$ if and only if $f = g$,

(c) $\mathrm{d}(f,g) = \mathrm{d}(g,f)$, and

(d) the *Triangle Inequality*: $\mathrm{d}(f,h) \le \mathrm{d}(f,g) + \mathrm{d}(g,h)$.

In this case, X is a called a *metric space*. The value $\mathrm{d}(f,g)$ is the *distance* from f to g.

If we need to explicitly identify the metric we write "let X be a metric space with metric d" or "let (X, d) be a metric space."

A metric space need not be a vector space, although this will be true of most of the metric spaces encountered in this volume.

Once we have a notion of distance, we have a corresponding notion of convergence.

Definition A.2 (Convergent and Cauchy Sequences). Let X be a metric space with metric d, and let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of elements of X.

(a) We say that $\{f_n\}_{n\in\mathbb{N}}$ *converges* to $f \in X$ if $\lim\limits_{n\to\infty} \mathrm{d}(f_n, f) = 0$, i.e., if

$$\forall\, \varepsilon > 0, \quad \exists\, N > 0, \quad \forall\, n \ge N, \quad \mathrm{d}(f_n, f) < \varepsilon.$$

In this case, we write $\lim\limits_{n\to\infty} f_n = f$ or $f_n \to f$.

(b) We say that $\{f_n\}_{n\in\mathbb{N}}$ is *Cauchy* if

$$\forall\, \varepsilon > 0, \quad \exists\, N > 0, \quad \forall\, m, n \ge N, \quad \mathrm{d}(f_m, f_n) < \varepsilon.$$

Exercise A.3. Let X be a metric space.

(a) Every convergent sequence in X is Cauchy.

(b) The limit of a convergent sequence is unique.

In general, however, a Cauchy sequence need not be convergent (Exercises A.63–A.61).

Definition A.4 (Complete Metric Space). If every Cauchy sequence in a metric space X has the property that it converges to an element of X, then X is said to be *complete*.

Beware that the term "complete" is heavily overused and has a number of distinct mathematical meanings.

Notation A.5. Let X be a metric space. Given $f \in X$ and $r > 0$, the *open ball* in X of radius r centered at f is

$$B_r(f) \;=\; \bigl\{g \in X : \mathrm{d}(f,g) < r\bigr\}. \tag{A.1}$$

A.3 Norms and Seminorms

A norm provides a notion of the length of a vector in a vector space.

Definition A.6 (Seminorms and Norms). Let X be a vector space over the field $\mathbb{C}$ of complex scalars. A *seminorm* on X is a function $\|\cdot\|\colon X \to \mathbb{R}$ such that for all f, $g \in X$ and all scalars $c \in \mathbb{C}$ we have:

(a) $\|f\| \ge 0$,

(b) $\|cf\| = |c|\,\|f\|$, and

(c) the *Triangle Inequality*: $\|f+g\| \le \|f\| + \|g\|$.

A seminorm is a *norm* if we also have:

(d) $\|f\| = 0$ if and only if $f = 0$.

A vector space X together with a norm $\|\cdot\|$ is called a *normed linear space* or simply a *normed space*. If the norm is not clear from context, we may write $(X, \|\cdot\|)$ to denote that $\|\cdot\|$ is the norm on X.

If S is a subspace of a normed space X, then S is itself a normed space with respect to the norm on X (restricted to S).

Exercise A.7. If X is a normed space, then $\mathrm{d}(f,g) = \|f-g\|$ defines a metric on X, called the *induced metric*.

The Schwartz space (Definition 1.90) is an example of a metric space whose metric is not induced from any norm; another is ℓ^p with $0 < p < 1$ (see Exercise A.18).

Exercise A.8. Show that if X is a normed linear space, then the following statements hold.

(a) Reverse Triangle Inequality: $\big|\|f\| - \|g\|\big| \le \|f - g\|$.

(b) Continuity of the norm: $f_n \to f \implies \|f_n\| \to \|f\|$.

(c) Continuity of vector addition: $f_n \to f$ and $g_n \to g \implies f_n + g_n \to f + g$.

(d) Continuity of scalar multiplication: $f_n \to f$ and $\alpha_n \to \alpha \implies \alpha_n f_n \to \alpha f$.

(e) Boundedness of convergent sequences: If $\{f_n\}_{n\in\mathbb{N}}$ converges then we have $\sup \|f_n\| < \infty$.

(f) Boundedness of Cauchy sequences: If $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy then we have $\sup \|f_n\| < \infty$.

Definition A.9 (Banach Space). A normed linear space X is called a *Banach space* if it is complete, if every Cauchy sequence is convergent.

Thus, the terms "Banach space" and "complete normed space" are interchangeable.

An important fact that we will assume without proof is that the complex plane $\mathbb{C}$ under absolute value is a Banach space.

A.3.1 Infinite Series in Normed Spaces

Since a normed space has both an operation of vector addition and a notion of convergence, we can consider infinite series.

Definition A.10 (Convergent Series). Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in a normed linear space X. Then the series $\sum_{n=1}^\infty f_n$ *converges and equals* $f \in X$ if the *partial sums* $s_N = \sum_{n=1}^N f_n$ converge to f, i.e., if

$$\lim_{N\to\infty} \|f - s_N\| \;=\; \lim_{N\to\infty} \Big\| f - \sum_{n=1}^N f_n \Big\| \;=\; 0.$$

Note that the ordering of a series may be important! If we reorder a series, or in other words consider a new series $\sum_{n=1}^\infty f_{\sigma(n)}$ where $\sigma\colon \mathbb{N} \to \mathbb{N}$ is a bijection, there is no guarantee that this reordered series will still converge. These issues are addressed in more detail in Section A.11.

Definition A.11 (Absolutely Convergent Series). Let X be a normed space and let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of elements of X. If

$$\sum_{n=1}^\infty \|f_n\| \;<\; \infty,$$

then we say that the series $\sum_{n=1}^\infty f_n$ is *absolutely convergent* in X.

The definition of absolute convergence does not require that the series $\sum f_n$ converge in X. This will always be the case if X is a Banach space, and indeed this property is an equivalent characterization of completeness.

Exercise A.12. Let X be a normed space. Prove that X is a Banach space if and only if every absolutely convergent series in X converges in X.

A.3.2 Convexity

Definition A.13 (Convex Set). If X is a vector space and $K \subseteq X$, then K is *convex* if

$$x, y \in K,\ 0 \le t \le 1 \implies tx + (1-t)y \in K.$$

Thus, the entire line segment between x and y is contained in K (including the midpoint $\frac{1}{2}x + \frac{1}{2}y$ in particular).

Every subspace of a vector space is convex by definition. The fact that balls in a normed space are convex is an important property.

Exercise A.14. Show that if X is a normed linear space, then each open ball $B_r(f)$ in X is convex.

Additional Problems

A.1. If $\{f_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in a normed space X and there exists a subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ that converges to $f \in X$, then $f_n \to f$.

A.2. If $\{f_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in a normed space X, then there exists a subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ such that $\|f_{n_{k+1}} - f_{n_k}\| < 2^{-k}$ for all $k \in \mathbb{N}$.

A.3. Let X be a normed space. Show that if $f_n \in X$ satisfy $\|f_{n+1} - f_n\| < 2^{-n}$ for every n, then $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy.

A.4. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in a normed space X, and let $f \in X$ be fixed. Suppose that every subsequence $\{g_n\}_{n\in\mathbb{N}}$ of $\{f_n\}_{n\in\mathbb{N}}$ has a subsequence $\{h_n\}_{n\in\mathbb{N}}$ of $\{g_n\}_{n\in\mathbb{N}}$ such that $h_n \to f$. Show that $f_n \to f$.

A.5. Let X be a normed space. Extend the definition of convergence to families indexed by a real parameter by declaring that if $f \in X$ and $f_t \in X$ for $t \in \mathbb{R}$, then $f_t \to f$ as $t \to 0$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $\|f - f_t\| < \varepsilon$ whenever $|t| < \delta$. Show that $f_t \to f$ as $t \to 0$ if and only if $f_{t_k} \to f$ for every sequence of real numbers $\{t_k\}_{k\in\mathbb{N}}$ such that $t_k \to 0$.

A.4 Examples of Banach Spaces: ℓ^p, C_b, C_0, C_b^m

In this section we give a few examples of Banach and other spaces.

A.4.1 The ℓ^p Spaces

Definition A.15. Let I be a finite or countably infinite index sequence.

(a) If $0 < p < \infty$, then $\ell^p(I)$ consists of all sequences of scalars $x = (x_k)_{k\in I}$ such that

$$\|x\|_p \;=\; \|(x_k)_{k\in I}\|_p \;=\; \biggl(\sum_{k\in I} |x_k|^p\biggr)^{1/p} \;<\; \infty.$$

(b) For $p = \infty$, the space $\ell^\infty(I)$ consists of all sequences of scalars $x = (x_k)_{k\in I}$ such that

$$\|x\|_\infty \;=\; \|(x_k)_{k\in I}\|_\infty \;=\; \sup_{k\in I} |x_k| \;<\; \infty.$$

If $I = \mathbb{N}$, then we write ℓ^p instead of $\ell^p(\mathbb{N})$.

If $I = \{1, \dots, d\}$, then $\ell^p(I) = \mathbb{C}^d$, and in this case we refer to $\ell^p(I)$ as "$\mathbb{C}^d$ under the ℓ^p norm." The ℓ^2 norm on $\mathbb{C}^d$ is called the *Euclidean norm.*

Now we prove a fundamental inequality for the ℓ^p spaces.

Theorem A.16 (Hölder's Inequality). *Let I be a finite or countable index set. Given $1 \le p \le \infty$, if $x = (x_k)_{k\in I} \in \ell^p(I)$ and $y = (y_k)_{k\in I} \in \ell^{p'}(I)$, then $xy = (x_k y_k)_{k\in I} \in \ell^1(I)$ and*

$$\|xy\|_1 \le \|x\|_p \, \|y\|_{p'}.$$

For $1 < p < \infty$, this inequality is

$$\sum_{k\in I} |x_k y_k| \;\le\; \biggl(\sum_{k\in I} |x_k|^p\biggr)^{1/p} \biggl(\sum_{k\in I} |y_k|^{p'}\biggr)^{1/p'}.$$

Proof. The cases $p = 1$ and $p = \infty$ are straightforward exercises. Assume $1 < p < \infty$. The key to the proof is a special case of an inequality due to Young for continuous, strictly increasing functions. Namely, since x^{p-1} is continuous and strictly increasing and its inverse function is $y^{\frac{1}{p-1}}$, we have for all $a,\, b \ge 0$ that

$$ab \;\le\; \int_0^a x^{p-1}\,dx + \int_0^b y^{\frac{1}{p-1}}\,dy \;=\; \frac{a^p}{p} + \frac{b^{p'}}{p'}$$

(see the "proof by picture" in Figure A.1, or Problem A.10).

Consequently, if $x \in \ell^p(I)$ and $y \in \ell^{p'}(I)$ satisfy $\|x\|_p = 1 = \|y\|_{p'}$, then

$$\|xy\|_1 \;=\; \sum_{k\in I} |x_k y_k| \;\le\; \sum_{k\in I} \biggl(\frac{|x_k|^p}{p} + \frac{|y_k|^{p'}}{p'}\biggr) \;=\; \frac{1}{p} + \frac{1}{p'} \;=\; 1. \tag{A.2}$$

For general nonzero x, y, we apply (A.2) to the normalized vectors $x/\|x\|_p$ and $y/\|y\|_{p'}$ to obtain

$$\frac{\|xy\|_1}{\|x\|_p \, \|y\|_{p'}} \;=\; \biggl\| \frac{x}{\|x\|_p} \, \frac{y}{\|y\|_{p'}} \biggr\|_1 \;\le\; 1. \qquad \square$$

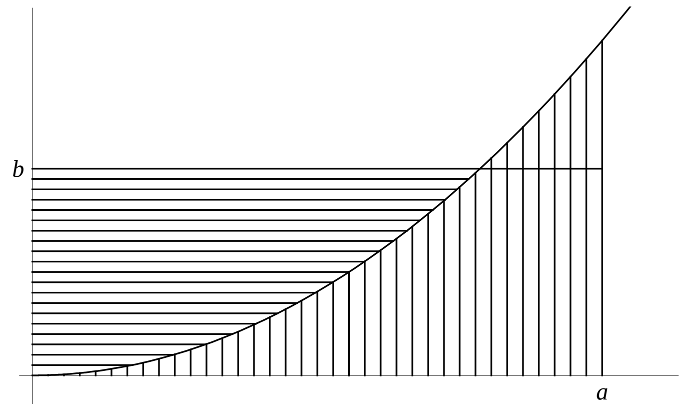

Fig. A.1. Illustration of Young's Inequality. Area of the vertically hatched region is $\int_0^a x^{p-1}\,dx$; area of the horizontally hatched region is $\int_0^b y^{\frac{1}{p-1}}\,dy$; area of the rectangle is ab.

The next exercise shows that if $p \geq 1$ then $\|\cdot\|_p$ is a norm on $\ell^p(I)$. The Triangle Inequality on ℓ^p (often called *Minkowski's Inequality*) is easy to prove for $p = 1$ and $p = \infty$, but more difficult for $1 < p < \infty$. A hint for using Hölder's Inequality to prove Minkowski's Inequality is given in the solutions section at the end of the text.

Exercise A.17. Let I be a finite or countable index set. Show that if $1 \leq p \leq \infty$, then $\|\cdot\|_p$ is a norm on $\ell^p(I)$, and $\ell^p(I)$ is a Banach space with respect to this norm.

On the other hand, if $p < 1$ then $\|\cdot\|_p$ fails the Triangle Inequality and hence is not a norm. Still, we can modify the distance function so that ℓ^p is a complete metric space, though this metric is not induced from any norm.

Exercise A.18. Let I be a finite or countably infinite index set. Show that if $0 < p < 1$, then $\|x+y\|_p^p \leq \|x\|_p^p + \|y\|_p^p$. Consequently, $\ell^p(I)$ is a vector space and $\mathrm{d}(x,y) = \|x-y\|_p^p$ is a metric on $\ell^p(I)$. Show that $\ell^p(I)$ is complete with respect to this metric. However, if I contains more than one element, then the unit ball $B_1(0)$ is not convex, and hence this metric is not induced from any norm (compare Exercise A.14).

We can also define $\ell^p(I)$ when the index set I is uncountable. In this case, for $p < \infty$ we define $\ell^p(I)$ to be the space of all sequences $x = (x_k)_{k\in I}$ with at most countably many terms nonzero such that $\sum |x_k|^p < \infty$. With this definition, $\ell^p(I)$ is again a Banach space.

A.4.2 Some Spaces of Continuous and Differentiable Functions

We now give some examples of normed spaces of functions.

Definition A.19. The *support* of a function $f\colon \mathbb{R} \to \mathbb{C}$ is the closure of the set of points where f is nonzero:

$$\operatorname{supp}(f) \;=\; \overline{\{x \in \mathbb{R} : f(x) \neq 0\}}.$$

Since the support of a function is a closed set, a function on $\mathbb{R}$ has compact support if and only if it is zero outside of a finite interval.

Exercise A.20. Let $C_b(\mathbb{R})$ denote the space of continuous, bounded functions $f\colon \mathbb{R} \to \mathbb{C}$. Show that $C_b(\mathbb{R})$ is a Banach space with respect to the *uniform norm*

$$\|f\|_\infty \;=\; \sup_{t\in\mathbb{R}} |f(t)|.$$

Show that the subspace

$$C_0(\mathbb{R}) \;=\; \bigl\{f \in C_b(\mathbb{R}) : \lim_{|t|\to\infty} f(t) = 0\bigr\}$$

is also a Banach space with respect to the uniform norm, but the subspace

$$C_c(\mathbb{R}) \;=\; \bigl\{f \in C_b(\mathbb{R}) : \operatorname{supp}(f) \text{ is compact}\bigr\} \tag{A.3}$$

is not complete with respect to the uniform norm.

Beware, some authors use the symbols C_0 to denote the space that we refer to as C_c.

Exercise A.20 has an extension to m-times differentiable functions, as follows.

Exercise A.21. Let $C_b^m(\mathbb{R})$ be the space of all m-times differentiable functions on $\mathbb{R}$ each of whose derivatives is bounded and continuous, i.e.,

$$C_b^m(\mathbb{R}) \;=\; \bigl\{f \in C_b(\mathbb{R}) : f, f', \dots, f^{(m)} \in C_b(\mathbb{R})\bigr\}.$$

Show that $C_b^m(\mathbb{R})$ is a Banach space with respect to the norm

$$\|f\|_{C_b^m} \;=\; \|f\|_\infty + \|f'\|_\infty + \cdots + \|f^{(m)}\|_\infty,$$

and

$$C_0^m(\mathbb{R}) \;=\; \bigl\{f \in C_0(\mathbb{R}) \;:\; f, f', \dots, f^{(m)} \in C_0(\mathbb{R})\bigr\}$$

is a subspace of $C_b^m(\mathbb{R})$ that is also a Banach space with respect to the same norm. However,

$$C_c^m(\mathbb{R}) \;=\; \bigl\{f \in C_c(\mathbb{R}) \;:\; f, f', \dots, f^{(m)} \in C_c(\mathbb{R})\bigr\}$$

is not a Banach space with respect to this norm.

Although they are not normed spaces, it is often important to consider the space of functions that are continuous or m-times differentiable but not bounded. We denote these by:

$$\begin{aligned} C(\mathbb{R}) &= \big\{f\colon \mathbb{R}\to\mathbb{C} \ : \ f \text{ is continuous}\big\}, \\ C^m(\mathbb{R}) &= \big\{f\in C(\mathbb{R}) \ : \ f, f', \dots, f^{(m)} \in C(\mathbb{R})\big\}. \end{aligned}$$

Additionally, we sometimes need to consider spaces of infinitely differentiable functions, including the following:

$$\begin{aligned} C^\infty(\mathbb{R}) &= \big\{f\in C(\mathbb{R}) \ : \ f, f', \dots \in C(\mathbb{R})\big\}, \\ C_b^\infty(\mathbb{R}) &= \big\{f\in C_b(\mathbb{R}) \ : \ f, f', \dots \in C_b(\mathbb{R})\big\}. \\ C_0^\infty(\mathbb{R}) &= \big\{f\in C_0(\mathbb{R}) \ : \ f, f', \dots \in C_0(\mathbb{R})\big\}. \\ C_c^\infty(\mathbb{R}) &= \big\{f\in C_c(\mathbb{R}) \ : \ f, f', \dots \in C_c(\mathbb{R})\big\}. \end{aligned}$$

The space $C_c^\infty(\mathbb{R})$ will be especially important to us in Chapter 3 and in Appendix E. Although not a normed space, it is topological vector space, and is often denoted by $\mathcal{D}(\mathbb{R}) = C_c^\infty(\mathbb{R})$.

Additional Problems

A.6. Fix $0 < p \le \infty$. Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence of vectors in $\ell^p(I)$, and x a vector in $\ell^p(I)$. Write the components of x_n and x as $x_n = (x_n(1), x_n(2), \dots)$ and $x = (x(1), x(2), \dots)$, and prove the following statements.

(a) If $x_n \to x$ in $\ell^p(I)$, then x_n *converges componentwise* to x, i.e., for each fixed k we have $\lim_{n\to\infty} x_n(k) = x(k)$.

(b) If I is finite then componentwise convergence implies convergence with respect to the norm $\|\cdot\|_p$.

(c) If I is infinite then componentwise convergence need not imply convergence in the norm of $\ell_p(I)$.

A.7. Show that if $1 \le p < q \le \infty$, then $\ell^p \subsetneq \ell^q$, and $\|x\|_q \le \|x\|_p$ for all $x \in \ell^p$.

A.8. Show that if $x \in \ell^q(I)$ for some finite q then $\|x\|_p \to \|x\|_\infty$ as $p\to\infty$, but this can fail if $x \notin \ell^q(I)$ for any finite q.

A.9. Let I be a finite or countable index set, and let $w\colon I \to (0,\infty)$ be fixed. Given a sequence of scalars $x = (x_k)_{k\in I}$, set

$$\|x\|_{p,w} \ = \ \begin{cases} \displaystyle\biggl(\sum_{k\in I} |x_k|^p\, w(k)^p\biggr)^{1/p}, & 0<p<\infty, \\ \displaystyle\sup_{k\in I}\, |x_k|\, w(k), & p=\infty, \end{cases}$$

and define the *weighted* ℓ^p *space* $\ell^p_w(I) = \big\{x : \|x\|_{p,w} < \infty\big\}$. Show that $\ell^p_w(I)$ is a Banach space for each $1 \le p \le \infty$.

A.10. (a) Show that if $0 < \theta < 1$, then $t^\theta \le \theta t + (1-\theta)$ for $t > 0$, with equality if and only if $t = 1$.

(b) Suppose that $1 < p < \infty$ and $a, b \ge 0$. Apply part (a) with $t = a^p b^{-p'}$ and $\theta = 1/p$ to show that $ab \le a^p/p + b^{p'}/p'$, with equality if and only if $b = a^{p-1}$.

A.11. Show that equality holds in Hölder's Inequality (Theorem A.16) if and only if there exist scalars α, β, not both zero, such that $\alpha\,|x_k|^p = \beta\,|y_k|^{p'}$ for each $k \in I$.

A.5 Inner Products

While a norm on a vector space provides a notion of the length of a vector, an inner product provides us with a notion of the angle between vectors.

Definition A.22 (Semi-Inner Product, Inner Product). If H is a vector space over the complex field $\mathbb{C}$, then a *semi-inner product* on H is a function $\langle\cdot,\cdot\rangle\colon H \times H \to \mathbb{C}$ such that for all f, g, $h \in H$ and scalars α, $\beta \in \mathbb{C}$ we have:

(a) $\langle f, f\rangle \ge 0$,

(b) $\langle f, g\rangle = \overline{\langle g, f\rangle}$,

(c) Linearity in the first variable: $\langle \alpha f + \beta g, h\rangle = \alpha\langle f, h\rangle + \beta\langle g, h\rangle$.

If a semi-inner product $\langle\cdot,\cdot\rangle$ also satisfies:

(d) $\langle f, f\rangle = 0$ if and only if $f = 0$,

then it is called an *inner product* on H. In this case, H is called a *inner product space* or a *pre-Hilbert space*.

There are many different standard notations for semi-inner products, including $[f, g]$, (f, g), or $\langle f|g\rangle$, in addition to our preferred notation $\langle f, g\rangle$.

Exercise A.23. If $\langle\cdot,\cdot\rangle$ is a semi-inner product on a vector space H, show that the following statements hold.

(a) Antilinearity in the second variable: $\langle f, \alpha g + \beta h\rangle = \bar{\alpha}\langle f, g\rangle + \bar{\beta}\langle f, h\rangle$.

(b) $\langle f, 0\rangle = 0 = \langle 0, f\rangle$.

(c) $\langle 0, 0\rangle = 0$.

A function of two variables that is linear in the first variable and antilinear (also called *conjugate linear*) in the second variable *conjugate linear function* is referred to as a *sesquilinear form.* Thus each semi-inner product $\langle\cdot,\cdot\rangle$ is an example of a sesquilinear form.[1] Sometimes an inner product is required to be antilinear in the first variable and linear in the second; this is common in the physics literature.

[1] The prefix "sesqui-" means "one and a half."

Every subspace S of an inner product space H is itself an inner product space (using the inner product on H restricted to S).

The next exercise gives the prototypical example of an inner product.

Exercise A.24. Given an index set I and $x = (x_k)_{k\in I}$, $(y_k)_{k\in I} \in \ell^2(I)$, set

$$\langle x, y\rangle \;=\; \sum_{k\in I} x_k\, \overline{y}_k.$$

Show that $\langle\cdot,\cdot\rangle$ is an inner product on $\ell^2(I)$.

Our next goal is to show that every semi-inner product induces a seminorm on H, and every inner product induces a norm.

Notation A.25. If $\langle\cdot,\cdot\rangle$ is a semi-inner product on a vector space H, then we write

$$\|f\| \;=\; \langle f, f\rangle^{1/2}, \quad f \in H.$$

Prejudicing the issue, we refer to $\|\cdot\|$ as the *seminorm induced by* $\langle\cdot,\cdot\rangle$. If $\langle\cdot,\cdot\rangle$ is an inner product, then we call $\|\cdot\|$ the *norm induced by* $\langle\cdot,\cdot\rangle$.

Before showing that $\|\cdot\|$ actually is a seminorm or norm, we derive some of its basic properties.

Exercise A.26. Given a semi-inner product $\langle\cdot,\cdot\rangle$ on a vector space H, show that the following statements hold for all f, $g \in H$.

(a) Polar Identity: $\|f+g\|^2 = \|f\|^2 + 2\,\mathrm{Re}\langle f, g\rangle + \|g\|^2$.

(b) Parallelogram Law: $\|f+g\|^2 + \|f-g\|^2 = 2\big(\|f\|^2 + \|g\|^2\big)$.

Now we prove an important inequality; this should be compared to Hölder's Inequality (Theorem A.16) for $p = 2$. This inequality is variously known as the *Schwarz*, *Cauchy–Schwarz*, or *Cauchy–Bunyakowski–Schwarz Inequality*.

Theorem A.27 (Cauchy–Bunyakowski–Schwarz Inequality). *If $\langle\cdot,\cdot\rangle$ is a semi-inner product on a vector space H, then*

$$\forall\, f, g, \in H, \quad |\langle f, g\rangle| \;\le\; \|f\|\,\|g\|.$$

Proof. If $f = 0$ or $g = 0$ then there is nothing to prove, so suppose both are nonzero. Write $\langle f, g\rangle = \alpha\,|\langle f, g\rangle|$ where $\alpha \in \mathbb{C}$ and $|\alpha| = 1$. Then for $t \in \mathbb{R}$ we have by the Polar Identity that

$$\begin{aligned} 0 \;\le\; \|f - \alpha t g\|^2 \;&=\; \|f\|^2 - 2\,\mathrm{Re}\bar{\alpha} t\langle f, g\rangle + t^2\,\|g\|^2 \\ &=\; \|f\|^2 - 2t\,|\langle f, g\rangle| + t^2\,\|g\|^2. \end{aligned}$$

This is a real-valued quadratic polynomial in the variable t. In order for it to be nonnegative, it can have at most one real root. This requires that the discriminant be at most zero, so $\big(-2\,|\langle f, g\rangle|\big)^2 - 4\,\|f\|^2\,\|g\|^2 \le 0$. The desired inequality then follows upon rearranging. □

When $\langle f, g \rangle = 0$, we say that f and g are *orthogonal*. More details on orthogonality appear in Section A.12.

Finally, the Cauchy–Bunyakowski–Schwarz Inequality can be used to show that $\|\cdot\|$ is indeed a seminorm or norm on H.

Exercise A.28. Given a semi-inner product $\langle \cdot, \cdot \rangle$ on a vector space H, show that $\|\cdot\|$ is a seminorm on H, and if $\langle \cdot, \cdot \rangle$ is an inner product then $\|\cdot\|$ is a norm on H.

Thus, all inner product spaces are normed linear spaces. The following exercise shows that there are normed space whose norm is not induced from any inner product on the space.

Exercise A.29. Let I be an index set containing at least two elements. Show that $\|\cdot\|_p$ does not satisfy the Parallelogram Law if $1 \le p \le \infty$ and $p \ne 2$. Therefore, there is no inner product on $\ell^p(I)$ whose induced norm is $\|\cdot\|_p$.

On the other hand, it is certainly possible for $\ell^p(I)$ to be an inner product space with respect to *some* inner product. For example, if I is finite or if $p < 2$, then $\ell^p(I) \subseteq \ell^2(I)$, so in this case $\ell^p(I)$ is an inner product space with respect to the inner product on $\ell^2(I)$ restricted to the subspace $\ell^p(I)$. However, the norm induced from this inner product is $\|\cdot\|_2$ and not $\|\cdot\|_p$. Further, if I is infinite then $\ell^p(I)$ is not complete with respect to the induced norm $\|\cdot\|_2$, whereas it is complete with respect to the norm $\|\cdot\|_p$.

Definition A.30 (Hilbert Space). An inner product space H is called a *Hilbert space* if it is complete with respect to the induced norm, i.e., if every Cauchy sequence is convergent.

Thus, a Hilbert space is an inner product space that is a Banach space with respect to the induced norm. In particular, $\ell^2(I)$ is a Hilbert space.

Additional Problems

A.12. Show that $\langle \cdot, \cdot \rangle$ is an inner product on $\mathbb{C}^d$ if and only if there exists a positive definite matrix A such that $\langle x, y \rangle = Ax \cdot y$, where $x \cdot y = x_1 \overline{y}_1 + \cdots + x_d \overline{y}_d$ denotes the usual dot product on $\mathbb{C}^d$.

A.13. Continuity of the inner product: If H is an inner product space and $f_n \to f$, $g_n \to g$ in H, then $\langle f_n, g_n \rangle \to \langle f, g \rangle$.

A.14. Let H be an inner product space. Show that if the series $\sum_{n=1}^{\infty} f_n$ converges in H, then for any $g \in H$,

$$\left\langle \sum_{n=1}^{\infty} f_n,\, g \right\rangle \;=\; \sum_{n=1}^{\infty} \langle f_n, g \rangle.$$

Note that this is not merely a consequence of the linearity of the inner product in the first variable — the continuity of the inner product is also needed.

A.15. Show that equality holds in the Cauchy–Bunyakowski–Schwarz Inequality if and only if there exist scalars α, $\beta \in \mathbb{C}$, not both zero, such that $\|\alpha f + \beta g\| = 0$. In particular, if $\|\cdot\|$ is a norm, then either $f = cg$ or $g = cf$ for some scalar c.

A.16. Justify the following statement: The angle between two vectors f, g in a Hilbert space H is the value of θ that satisfies $\mathrm{Re}\langle f, g\rangle = \|f\|\,\|g\|\cos\theta$.

A.6 Topology

Now we consider topologies, especially on metric and normed spaces. General background references on topology include Munkres [Mun75] and Singer and Thorpe [ST76].

Definition A.31 (Topology). A *topology* on a set X is a family $\mathcal{T}$ of subsets of X such that the following statements hold.

(a) $\emptyset$, $X \in \mathcal{T}$.

(b) Closure under arbitrary unions: If I is any index set and $U_i \in \mathcal{T}$ for $i \in I$, then $\cup_{i\in I} U_i \in \mathcal{T}$.

(c) Closure under finite intersections: If U, $V \in \mathcal{T}$, then $U \cap V \in \mathcal{T}$.

If these hold then $\mathcal{T}$ is called a *topology* on X and X is called a *topological space*. The elements of $\mathcal{T}$ are called the *open subsets* of X. The complements of the open subsets are the *closed subsets* of X. A *neighborhood* of a point $x \in X$ is any set A such that there exists an open set U with $x \in U \subseteq A$. In particular, an *open neighborhood* of x is any open set U that contains x. If the topology on X is not clear from context, we may write $(X, \mathcal{T})$ to denote that $\mathcal{T}$ is the topology on X.

The following is a convenient criterion for testing for openness.

Exercise A.32. Let X be a topological space. Prove that $V \subseteq X$ is open if and only if

$$\forall\, x \in V, \quad \exists \text{ open } U \subseteq V \text{ such that } x \in U.$$

If a space X has two topologies $\mathcal{T}_1$, $\mathcal{T}_2$ and if $\mathcal{T}_1 \subseteq \mathcal{T}_2$, i.e., every set that is open with respect to $\mathcal{T}_1$ is also open with respect to $\mathcal{T}_2$, then we say that $\mathcal{T}_1$ *is weaker than* $\mathcal{T}_2$ and $\mathcal{T}_2$ *is stronger than* $\mathcal{T}_1$. These terms should not be confused with the *weak topology* or the *strong topology* on a space. The strong topology on a normed space is defined below (see Definition A.34) and the weak topology is defined in Example E.7.

The most familiar topologies are those associated with metric spaces.

Exercise A.33. Let (X, d) be a metric space. Declare a subset $U \subseteq X$ to be *open* if

$$\forall\, f \in U, \quad \exists\, r > 0 \text{ such that } B_r(f) \subseteq U,$$

where $B_r(f)$ is the open ball centered at f with radius r. Let $\mathcal{T}$ be the collection of all subsets of X that are open according to this definition. Show that $\mathcal{T}$ is a topology on X.

Thus every metric space has a natural topology associated with it, and consequently so does every normed space.

Definition A.34. (a) If (X, d) is a metric space, then the topology $\mathcal{T}$ defined in Exercise A.33 is called the *topology on X induced from the metric* d, or simply the *induced topology* on!X.

(b) If $(X, \|\cdot\|)$ is a normed linear space, then the topology $\mathcal{T}$ induced from the metric $\mathrm{d}(f, g) = \|f - g\|$ is called the *norm topology*, the *strong topology*, or the *topology induced from* $\|\cdot\|$.

(c) Let $(X, \mathcal{T})$ be a topological space. If there exists a metric d on X whose induced topology is exactly $\mathcal{T}$, then the topology $\mathcal{T}$ on X is said to be *metrizable*.

(d) Let X be a vector space with a topology $\mathcal{T}$. If there exists a norm $\|\cdot\|$ on X whose induced topology is exactly $\mathcal{T}$, then the topology $\mathcal{T}$ on X is said to be *normable*.

The Schwartz space $\mathcal{S}(\mathbb{R})$ and the space $C^\infty(\mathbb{R})$ are important examples of topological spaces that are metrizable but not normable (see Example E.28).

The topology induced by a metric has the following special property.

Definition A.35. A topological space X is *Hausdorff* if

$$\forall\, x \neq y \in X, \quad \exists \text{ disjoint } U, V \in \mathcal{T} \text{ such that } x \in U,\ y \in V.$$

Exercise A.36. Every metric space is Hausdorff.

Every subset of a topological space X inherits a topology from X.

Exercise A.37. Let X be a topological space. Given $Y \subseteq X$, show that

$$\mathcal{T}_Y \;=\; \{U \cap Y : U \text{ is open in } X\}$$

is a topology on Y, called the *topology on Y relative to X*, the *topology on Y inherited from X*, the *topology on X restricted to Y*, etc.

One way to generate a topology on a set X is to begin with a collection of sets that we want to be open, and then to create a topology that includes those particular sets. There will be many such topologies in general, but the following exercise shows that there is a smallest topology that includes those chosen sets.

Exercise A.38. Let $\mathcal{E}$ be a collection of subsets of a set X. Show that

$$\mathcal{T}(\mathcal{E}) \;=\; \bigcap \bigl\{\mathcal{T} : \mathcal{T} \text{ is a topology on } X \text{ with } \mathcal{E} \subseteq \mathcal{T}\bigr\}$$

is a topology on X. We call $\mathcal{T}(\mathcal{E})$ the *topology generated by* $\mathcal{E}$. The collection $\mathcal{E}$ is sometimes called a *subbase* for the topology $\mathcal{T}(\mathcal{E})$.

Note that if $\mathcal{T}_1$, $\mathcal{T}_2$ are topologies, then $\mathcal{T}_1 \cap \mathcal{T}_2$ is not formed by intersecting the elements of $\mathcal{T}_1$ with those of $\mathcal{T}_2$. Rather, it is the collection of all sets that are common to both $\mathcal{T}_1$ and $\mathcal{T}_2$. Thus, if $\mathcal{T}$ is any topology that contains $\mathcal{E}$ then we will have $\mathcal{T}(\mathcal{E}) \subseteq \mathcal{T}$, which explains why $\mathcal{T}(\mathcal{E})$ is the *smallest* topology that contains $\mathcal{E}$. In particular, the topology induced by a metric d on a metric space X is the topology generated by the set of open balls in X.

We can characterize the generated topology $\mathcal{T}(\mathcal{E})$ as follows.

Exercise A.39. If $\mathcal{E}$ is a collection of subsets of a set X whose union is X, then $\mathcal{T}(\mathcal{E})$ is set of all unions of finite intersections of elements of $\mathcal{E}$:

$$\mathcal{T}(\mathcal{E}) \;=\; \biggl\{ \bigcup_{i\in I} \bigcap_{j=1}^{n} E_{ij} \;:\; I \text{ arbitrary},\; n \in \mathbb{N},\, E_{ij} \in \mathcal{E} \biggr\}.$$

A.6.1 Product Topologies

As an application of generated topologies, we show how to construct a natural topology on the Cartesian product of two topological spaces. A product topology on an infinite collection of topological spaces can also be defined but requires a little more care, see Definition E.44.

Definition A.40 (Product Topology). Let X and Y be topological spaces, and set

$$\mathcal{B} \;=\; \{U \times V : \text{open } U \subseteq X, \text{ open } V \subseteq Y\}.$$

The *product topology* on $X \times Y$ is the topology $\mathcal{T}(\mathcal{B})$ generated by $\mathcal{B}$.

Exercise A.41. (a) Show that $\mathcal{B}$ as given above is a *base* for $\mathcal{T}(\mathcal{B})$, which means that if $W \in \mathcal{T}(\mathcal{B})$, then there exist sets U_α open in X and V_α open in Y such that $W = \cup_\alpha (U_\alpha \times V_\alpha)$.

(b) Show that if $W \subseteq X \times Y$ is open with respect to the product topology, then for each $x \in X$ the restriction $W_x = \{y \in Y : (x,y) \in W\}$ is open in Y, and likewise for each $y \in Y$ the restriction $W^y = \{x \in X : (x,y) \in W\}$ is open in X.

A.7 Convergence and Continuity in Topological Spaces

In large part, the importance of topologies in this volume is that they provide notions of convergence and continuity. Indeed, a basic philosophy that we will expand upon in this section is that topologies and convergence criteria are equivalent. Thus, though many of the spaces that we will encounter are defined in terms of norms or families of seminorms (i.e., convergence criteria), this is equivalent to defining them in terms of a topology.

A.7.1 Convergence

In metric spaces, convergence is defined with respect to sequences indexed by the natural numbers (Definition A.2). In a general topological space, convergence must be formulated in terms of nets instead of countable sequences.

Definition A.42 (Directed Sets, Nets). A *directed set* is a set I together with a relation $\leq$ on I such that:

(a) $\leq$ is reflexive: $i \leq i$ for all $i \in I$,

(b) $\leq$ is transitive: $i \leq j$ and $j \leq k$ implies $i \leq k$, and

(c) for any $i, j \in I$, there exists $k \in I$ such that $i \leq k$ and $j \leq k$.

A *net* in a set X is a sequence $\{x_i\}_{i \in I}$ of elements of X indexed by a directed set $(I, \leq)$.

Remark A.43. By definition, a sequence $\{x_i\}_{i \in I}$ is shorthand for the function $x\colon I \to X$ defined by $x(i) = x_i$ for $i \in I$. In particular, unlike a set, a sequence allows repetitions of the x_i. Technically, we should be careful to distinguish between a sequence $\{x_i\}_{i \in I}$ and a set $\{x_i : i \in I\}$, but it is usually clear from context whether a sequence or a set is meant.

The set of natural numbers $I = \mathbb{N}$ under the usual ordering is one example of a directed set, and hence every ordinary sequence indexed by the natural numbers is a net. Another typical example is $I = \mathcal{P}(X)$, the power set of X, ordered by *reverse inclusion*, i.e.,

$$U \leq V \iff V \subseteq U.$$

Definition A.44 (Convergence of a Net). Let X be a topological space, let $\{x_i\}_{i \in I}$ be a net in X, and let $x \in X$ be given. Then we say that $\{x_i\}_{i \in I}$ *converges* to x (with respect to the directed set I), and write $x_i \to x$, if for any open neighborhood U of x there exists $i_0 \in I$ such that

$$i \geq i_0 \implies x_i \in U.$$

Next, we will define the notion of accumulation points of a subset of a generic topological space and see how this definition can be reformulated in terms of nets. We will also see that topologies induced from a metric have the advantage that we only need to use convergence of ordinary sequences indexed by $\mathbb{N}$ instead of general nets.

Definition A.45 (Accumulation Point). Let E be a subset of a topological space X. Then a point $x \in X$ is an *accumulation point* of E if every open neighborhood of x contains a point of E other than x itself, i.e.,

$$U \text{ open and } x \in U \implies E \cap (U \backslash \{x\}) \neq \emptyset.$$

Lemma A.46. *If E is a subset of a topological space X and $x \in X$, then the following statements are equivalent.*

(a) *x is an accumulation point of E.*

(b) *There exists a net $\{x_i\}_{i \in I}$ contained in $E \backslash \{x\}$ such that $x_i \to x$.*

If X is a metric space, then these statements are also equivalent to the following.

(c) *There exists a sequence $\{x_n\}_{n \in \mathbb{N}}$ contained in $E \backslash \{x\}$ such that $x_n \to x$.*

Proof. (a) $\Rightarrow$ (b). Assume that x is an accumulation point of E. Define

$$I = \bigl\{U \subseteq X : U \text{ is open and } x \in U\bigr\}.$$

Exercise: Show that I is a directed set when ordered by reverse inclusion.

For each $U \in I$, by definition of accumulation point there exists a point $x_U \in E \cap (U \backslash \{x\})$. Then $\{x_U\}_{U \in I}$ is a net in $E \backslash \{x\}$, and we claim that $x_U \to x$. To see this, fix any open neighborhood V of x. Set $U_0 = V$, and suppose that $U \geq U_0$. Then, by definition, $U \in I$ and $U \subseteq U_0$. Hence $x_U \in U \subseteq U_0 = V$. Therefore $x_U \to x$.

(b) $\Rightarrow$ (a). Suppose that $\{x_i\}_{i \in I}$ is a net in $E \backslash \{x\}$ and $x_i \to x$. Let U be any open neighborhood of x. Then there exists an i_0 such that $x_i \in U$ for all $i \geq i_0$. Since $x_i \neq x$, this implies that $x_i \in E \cap (U \backslash \{x\})$ for all $i \geq i_0$.

(a) $\Rightarrow$ (c), assuming X is metric. Suppose that x is an accumulation point of E. For each $n \in \mathbb{N}$, the open ball $B_{1/n}(x)$ is an open neighborhood of x, and hence there must exist some $x_n \in E \cap (B_{1/n}(x) \backslash \{x\})$. Therefore $\{x_n\}_{n \in \mathbb{N}}$ is a sequence in $E \backslash \{x\}$, and since $\mathrm{d}(x, x_n) < 1/n$, we have $x_n \to x$.

(c) $\Rightarrow$ (b), assuming X is metric. This follows from the fact that every countable sequence $\{x_n\}_{n \in \mathbb{N}}$ is a net. $\square$

We can now give an equivalent formulation of closed sets in terms of nets and accumulation points.

Exercise A.47. Given a subset E of a topological space, prove that the following statements are equivalent.

(a) E is closed, i.e., $X \backslash E$ is open.

(b) If x is an accumulation point of E, then $x \in E$.

(c) If $\{x_i\}_{i \in I}$ is a net in E and $x_i \to x \in X$, then $x \in E$.

If X is a metric space, show that these are also equivalent to the following statement.

(d) If $\{x_n\}_{n\in\mathbb{N}}$ is a sequence in E and $x_n \to x \in X$, then $x \in E$.

Now we can quantify the philosophy that topologies and convergence criteria are equivalent. For arbitrary topologies, this requires that we use convergence with respect to nets, but for topologies induced from a metric we are able to use convergence of ordinary sequences indexed by the natural numbers.

Exercise A.48. Given two topologies $\mathcal{T}_1$, $\mathcal{T}_2$ on a set X, prove that the following statements are equivalent.

(a) $\mathcal{T}_1 \subseteq \mathcal{T}_2$, i.e.,

$$U \text{ is open with respect to } \mathcal{T}_1 \implies U \text{ is open with respect to } \mathcal{T}_2.$$

(b) If $\{x_i\}_{i\in I}$ is a net in X and $x \in X$, then

$$x_i \to x \text{ with respect to } \mathcal{T}_2 \implies x_i \to x \text{ with respect to } \mathcal{T}_1.$$

If $\mathcal{T}_1$ is induced from a metric d_1 on X, and $\mathcal{T}_2$ is induced from a metric d_2 on X, show that these statements are also equivalent to the following.

(c) If $\{x_n\}_{n\in\mathbb{N}}$ is a sequence in X and $x \in X$, then

$$\lim_{n\to\infty} \mathrm{d}_2(x_n, x) = 0 \implies \lim_{n\to\infty} \mathrm{d}_1(x_n, x) = 0.$$

Interchanging the roles of $\mathcal{T}_1$ and $\mathcal{T}_2$ in Exercise A.48, we see that $\mathcal{T}_1 = \mathcal{T}_2$ if and only if $\mathcal{T}_1$ and $\mathcal{T}_2$ define exactly the same convergence criterion.

Example A.49. An example of a topological space where it is important to distinguish between convergence of ordinary sequences and convergence with respect to nets is the sequence space ℓ^1 under the weak topology. This topology will be defined precisely in Section E.6, but the important point for us at the moment is that it can be shown that if $\{x_n\}_{n\in\mathbb{N}}$ is a sequence in ℓ^1 and $x_n \to x$ with respect to the weak topology, then $x_n \to x$ in norm, i.e., $\|x - x_n\|_1 \to 0$ (see [Con90, Prop. V.5.2]). However, the weak topology on ℓ^1 is *not* the same as the topology induced by the norm $\|\cdot\|_1$. The moral is that when discussing convergence in a topological space that is not a metric space, it is important to consider nets instead of ordinary sequences.

A.7.2 Continuity

Our next goal is to reformulate continuity of a function in terms of convergence of nets or sequences. Recall that if $f\colon X \to Y$ and $V \subseteq Y$, then the preimage of V is $f^{-1}(V) = \{x \in X : f(x) \in V\}$.

Definition A.50 (Continuity). Let X, Y be topological spaces. Then a function $f\colon X \to Y$ is *continuous* if

$$V \text{ is open in } Y \implies f^{-1}(V) \text{ is open in } X.$$

We say that f is a *topological isomorphism* or a *homeomorphism* if f is a bijection and both f and f^{-1} are continuous.

It will be convenient to restate continuity in terms of continuity at a point.

Definition A.51 (Continuity at a Point). Let X, Y be topological spaces and let $x \in X$ be given. Then a function $f\colon X \to Y$ is *continuous at* x if for each open neighborhood V of $f(x)$ in Y, there exists an open neighborhood U of x in X such that $U \subseteq f^{-1}(V)$.

Exercise A.52. Prove that f is continuous if and only if f is continuous at each $x \in X$.

Now we can formulate continuity in terms of preservation of convergence of nets. For the case of a metric space, this reduces to preservation of convergence of sequences.

Lemma A.53. *If X, Y be topological spaces, $f\colon X \to Y$, and $x \in X$ are given, then the following statements are equivalent.*

(a) *f is continuous at x.*

(b) *For any net $\{x_i\}_{i \in I}$ in X,*

$$x_i \to x \text{ in } X \implies f(x_i) \to f(x) \text{ in } Y.$$

If X is a metric space, then these are also equivalent to the following.

(c) *For any sequence $\{x_n\}_{n \in \mathbb{N}}$ in X,*

$$x_n \to x \text{ in } X \implies f(x_n) \to f(x) \text{ in } Y.$$

Proof. (a) $\Rightarrow$ (b). Assume that f is continuous at $x \in X$. Let $\{x_i\}_{i \in I}$ be any net in X such that $x_i \to x$. Let V be any open neighborhood of $f(x)$. Then by definition of continuity at a point, there exists an open neighborhood U of x that is contained in $f^{-1}(V)$. Hence, by definition of $x_i \to x$, there exists an $i_0 \in I$ such that $x_i \in U \subseteq f^{-1}(V)$ for all $i \geq i_0$. Hence $f(x_i) \in V$ for all $i \geq i_0$, which means that $f(x_i) \to f(x)$.

(b) $\Rightarrow$ (a). Suppose that statement (a) fails, i.e., f is not continuous at $x \in X$. Then by definition there exists an open neighborhood V of $f(x)$ such that no open neighborhood U of x can be contained in $f^{-1}(V)$. Therefore each open neighborhood U of x must contain some point $x_U \in U \backslash f^{-1}(V)$. Now let

$$I = \big\{U : U \text{ is an open neighborhood of } x\big\}.$$

Then I is a directed set when ordered by reverse inclusion, so $\{x_U\}_{U \in I}$ is a net in X. Exercise: Show that $x_U \to x$. However, $f(x_U)$ does not converge to $f(x)$ because V is an open neighborhood of $f(x)$ but V contains no points $f(x_U)$. Hence statement (b) fails.

The remaining implications are exercises. $\square$

A.7.3 Equivalent Norms

Next we consider the equivalence of convergence criteria and topologies for the case of normed spaces.

Definition A.54. Suppose that X is a normed linear space with respect to a norm $\|\cdot\|_a$ and also with respect to another norm $\|\cdot\|_b$. Then we say that these norms are *equivalent* if there exist constants $C_1, C_2 > 0$ such that

$$\forall\, f \in X, \quad C_1\, \|f\|_a \;\le\; \|f\|_b \;\le\; C_2\, \|f\|_a. \tag{A.4}$$

We write $\|\cdot\|_a \asymp \|\cdot\|_b$ to denote that $\|\cdot\|_a$ and $\|\cdot\|_b$ are equivalent norms.

Theorem A.55. *Let $\|\cdot\|_a$ and $\|\cdot\|_b$ be two norms on a vector space X. Then the following statements are equivalent.*

(a) *$\|\cdot\|_a$ and $\|\cdot\|_b$ are equivalent norms.*

(b) *$\|\cdot\|_a$ and $\|\cdot\|_b$ induce the same topologies on X.*

(c) *$\|\cdot\|_a$ and $\|\cdot\|_b$ define the same convergence criterion. That is, if $\{x_n\}_{n\in\mathbb{N}}$ is a sequence in X and $x \in X$, then*

$$\lim_{n\to\infty} \|x - x_n\|_a = 0 \iff \lim_{n\to\infty} \|x - x_n\|_b = 0.$$

Proof. (b) $\Rightarrow$ (a). Assume that statement (b) holds. Let $B_r^a(x)$ and $B_r^b(x)$ denote the open balls of radius r centered at $x \in X$ with respect to $\|\cdot\|_a$ and $\|\cdot\|_b$, respectively. Since $B_1^a(0)$ is open with respect to $\|\cdot\|_a$, statement (b) implies that $B_1^a(0)$ is open with respect to $\|\cdot\|_b$. Therefore, since $0 \in B_1^a(0)$, there must exist some $r > 0$ such that $B_r^b(0) \subseteq B_1^a(0)$.

Now choose any $x \in X$ and any $\varepsilon > 0$. Then

$$\frac{(r-\varepsilon)}{\|x\|_b}\, x \;\in\; B_r^b(0) \;\subseteq\; B_1^a(0),$$

so

$$\left\| \frac{(r-\varepsilon)\, x}{\|x\|_b} \right\|_a \;<\; 1.$$

Rearranging, this implies $(r-\varepsilon)\, \|x\|_a < \|x\|_b$. Since this is true for every ε, we conclude that $r\, \|x\|_a \le \|x\|_b$.

A symmetric argument, interchanging the roles of the two norms, shows that there exists an $s > 0$ such that $\|x\|_b \le s\, \|x\|_a$ for every $x \in X$. Hence the two norms are equivalent. $\square$

Given any finite-dimensional vector space X, we can define many norms on X. In particular, the following norms are analogues of the ℓ^p norms defined in Section A.4.

Exercise A.56. Let $\mathcal{B} = \{x_1, \dots, x_d\}$ be any basis for a finite-dimensional vector space X, and let $x = \sum_{k=1}^{d} c_k(x)\, x_k$ denote the unique expansion of $x \in X$ with respect to this basis (the vector $[x]_{\mathcal{B}} = (c_1(x), \dots, c_d(x))$ is called the *coordinate vector* of x with respect to the basis $\mathcal{B}$). Show that

$$\|x\|_p \;=\; \begin{cases} \left(\displaystyle\sum_{k=1}^{d} |c_k(x)|^p\right)^{1/p}, & 1 \le p < \infty, \\ \displaystyle\max_k |c_k(x)|, & p = \infty, \end{cases}$$

are norms on X, and X is complete with respect to each of these norms. Note that $\|x\|_p$ is simply the ℓ^p norm of the coordinate vector $[x]_{\mathcal{B}}$.

It is not difficult to see that all of the norms defined in Exercise A.56 are equivalent. Although we will not prove it, it is an important fact that *all* norms on a finite-dimensional space are equivalent.

Theorem A.57. *If X is a finite-dimensional vector space, then any two norms on X are equivalent. In particular, if $\|\cdot\|$ is any norm on X and $\|\cdot\|_p$ is any one of the norms constructed in Exercise A.56, then $\|\cdot\| \asymp \|\cdot\|_p$.*

Additional Problems

A.17. (a) Let X be a Hausdorff topological space. Show that if a net $\{x_i\}_{i \in I}$ converges in X, then the limit is unique.

(b) Show that if X is not Hausdorff, then there exists a net $\{x_i\}_{i \in I}$ in X that has two distinct limits.

A.18. Let $\{x_i\}_{i \in I}$ be a net in a Hausdorff topological space X. Show that if $x_i \to x$ in X, then either:

(a) there exists an open neighborhood U of x and some $i_0 \in I$ such that $x_i = x$ for all $i \ge i_0$, or

(b) every open neighborhood U of x contains infinitely many distinct x_i, i.e., the set $\{x_i : i \in I \text{ and } x_i \in U\}$ is infinite.

A.19. Show that if (X, d_1) and (Y, d_2) are metric spaces, then

$$\mathrm{d}\big((f_1, g_1),\, (f_2, g_2)\big) \;=\; \mathrm{d}_1(f_1, f_2) + \mathrm{d}_2(g_1, g_2)$$

defines a metric on $X \times Y$ that induces the product topology. Conclude that convergence in $X \times Y$ is componentwise convergence, i.e., $(f_n, g_n) \to (f, g)$ in $X \times Y$ if and only if $f_n \to f$ in X and $g_n \to g$ in Y.

A.20. Let X be a vector space with a metric d. We say that the metric is *translation-invariant* if $\mathrm{d}(f + h, g + h) = \mathrm{d}(f, g)$ for every f, g, $h \in X$. Show that vector addition is continuous in this case, i.e., $(f, g) \to f + g$ is a continuous mapping of $X \times X$ into X.

A.21. Let X, Y be topological spaces. Let $\{(f_i, g_i)\}_{i\in I}$ be any net in $X \times Y$, and suppose $(f,g) \in X \times Y$. Show that $(f_i, g_i) \to (f,g)$ with respect to the product topology on $X \times Y$ if and only if $f_i \to f$ in X and $g_i \to g$ in Y.

A.8 Closed and Dense Sets

The smallest closed set that contains a given set is called its closure, defined precisely as follows.

Definition A.58. If E is a subset of a topological space X, then the *closure of* E, denoted $\overline{E}$, is the smallest closed set in X that contains E:

$$\overline{E} \;=\; \bigcap\{F \subseteq X : F \text{ is closed and } F \supseteq E\}.$$

If $\overline{E} = X$, then we say that E is *dense* in X.

Often it is more convenient to use the following equivalent form of the closure of a set.

Exercise A.59. Given a subset E of a topological space X, show that $\overline{E}$ is the union of E and all the accumulation points of E.

The typical method for showing that a subset of a metric space is dense is given in the next exercise.

Exercise A.60. Let X be a metric space, and let $E \subseteq X$ be given. Show that E is dense in X if and only if for each $f \in X$ there exist a sequence $\{f_n\}_{n\in\mathbb{N}}$ in E such that $f_n \to f$.

The next exercise characterizes the closed subspaces of a Banach space.

Exercise A.61. Let M be a subspace of a Banach space X. Then M is itself a Banach space (using the norm inherited from X) if and only if M is closed.

Every subspace of a finite-dimensional normed space is closed, but this need not be the case in infinite dimensions.

Exercise A.62. Show that every finite-dimensional subspace of a normed linear space is closed.

Exercise A.63. (a) Fix $1 \le p \le \infty$. Define

$$c_{00} \;=\; \big\{x = (x_1, \dots, x_N, 0, 0, \dots) : N > 0,\, x_1, \dots, x_N \in \mathbb{C}\big\}.$$

Prove that c_{00} is a proper dense subspace of ℓ^p for each $1 \le p < \infty$, and hence is not closed with respect to $\|\cdot\|_p$. The vectors in c_{00} are sometimes called *finite sequences* because they contain at most finitely many nonzero components.

(b) Define

$$c_0 \;=\; \Big\{x = (x_k)_{k=1}^{\infty} : \lim_{k\to\infty} x_k = 0\Big\}.$$

Prove that c_0 is a closed subspace of $\ell^\infty(\mathbb{N})$, and c_0 is the closure of c_{00} in ℓ^∞-norm.

(c) Show that $C_c(\mathbb{R})$ is a proper dense subspace of $C_0(\mathbb{R})$, and $C_0(\mathbb{R})$ is the closure of $C_c(\mathbb{R})$ in the uniform norm.

We now introduce a definition that in some sense distinguishes between "small" and "large" infinite-dimensional spaces.

Definition A.64. A topological space that contains a countable dense subset is said to be *separable*.

Exercise A.65. (a) Show that if I is a finite or countably infinite index set, then $\ell^p(I)$ is separable for $0 < p < \infty$.

(b) Show that if I is infinite, then $\ell^\infty(I)$ is not separable.

(c) Show that if I is uncountable, then $\ell^p(I)$ is not separable for any p.

Exercise A.66. Show that c_0 and $C_0(\mathbb{R})$ are separable.

A.9 Compact Sets

Now we briefly review the definition and properties of compact sets.

Definition A.67 (Compact Set). A subset K of a topological space X is *compact* if every cover of K by open sets has a finite subcover. That is, K is compact if it is the case that whenever

$$K \;\subseteq\; \bigcup_{i\in I} U_i,$$

where $\{U_i\}_{i\in I}$ is *any* collection of open subsets of X, there exist *finitely* many $i_1, \dots, i_N \in I$ such that

$$K \;\subseteq\; \bigcup_{k=1}^{N} U_{i_k}.$$

In a metric space, we can reformulate compactness in several equivalent ways. We need the following terminology.

Definition A.68. Let E be a subset of a metric space X.

(a) E is *sequentially compact* if every sequence $\{f_n\}_{n\in\mathbb{N}}$ of points of E contains a convergent subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ whose limit belongs to E.

(b) E is *complete* if every Cauchy sequence in E converges to a point in E.

(c) E is *totally bounded* if for every $r > 0$, we can cover E by finitely many open balls of radius r, i.e., there exist finitely many $x_1, \dots, x_N \in X$ such that

$$E \subseteq \bigcup_{k=1}^{N} B_r(x_k).$$

Similarly to Exercise A.61, if X is complete then a subset E is complete if and only if it is closed.

Theorem A.69. *If E is a subset of a metric space X, then the following statements are equivalent.*

(a) *E is compact.*

(b) *E is sequentially compact.*

(c) *E is complete and totally bounded.*

Proof. (a) $\Rightarrow$ (b). Suppose there exists a sequence $\{f_n\}_{n\in\mathbb{N}}$ in E that has no subsequence that converges to an element of E. Choose any $f \in E$. If every open ball centered at f contained infinitely many vectors f_n, then there would be a subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ that converged to f. Therefore, there must exist some open ball B_f centered at f that contains only finitely many f_n. But then $\{B_f\}_{f\in E}$ is an open cover of E that has no finite subcover. Hence E is not compact.

(b) $\Rightarrow$ (c). Suppose that E is sequentially compact. Exercise: Show that E is complete.

Suppose that E was not totally bounded. Then there is an $r > 0$ such that E cannot be covered by finitely many open balls of radius r centered at points of X. Choose any $f_1 \in E$. Since E cannot be covered by a single r-ball, E cannot be a subset of $B_r(f_1)$. Hence there exists a point $f_2 \in E \setminus B_r(f_1)$. In particular, we have $\mathrm{d}(f_2, f_1) \ge r$. But E cannot be covered by two r-balls, so there must exist an $f_3 \in E \setminus \big(B_r(f_1) \cup B_r(f_2)\big)$. In particular, we have both $\mathrm{d}(f_3, f_1) \ge r$ and $\mathrm{d}(f_3, f_2) \ge r$. Continuing in this way we obtain a sequence of points $\{f_n\}_{n\in\mathbb{N}}$ in E that has no convergent subsequence, which is a contradiction.

(c) $\Rightarrow$ (b). Assume that E is complete and totally bounded, and let $\{f_n\}_{n\in\mathbb{N}}$ be any sequence of points in E. Since E is covered by finitely many open balls of radius $\frac{1}{2}$, one of those balls must contain infinitely many f_n, say $\{f_n^{(1)}\}_{n\in\mathbb{N}}$. Then we have

$$\forall\, m, n \in \mathbb{N}, \quad \mathrm{d}\big(f_m^{(1)}, f_n^{(1)}\big) \;<\; 1.$$

Since E is covered by finitely many open balls of radius $\frac{1}{4}$, we can find a subsequence $\{f_n^{(2)}\}_{n\in\mathbb{N}}$ of $\{f_n^{(1)}\}_{n\in\mathbb{N}}$ such that

$$\forall\, m, n \in \mathbb{N}, \quad \mathrm{d}\big(f_m^{(2)}, f_n^{(2)}\big) \;<\; \frac{1}{2}.$$

Continuing by induction, for each $k > 1$ we find a subsequence $\{f_n^{(k)}\}_{n\in\mathbb{N}}$ of $\{f_n^{(k-1)}\}_{n\in\mathbb{N}}$ such that $\mathrm{d}\big(f_m^{(k)}, f_n^{(k)}\big) < \frac{1}{k}$ for all m, $n \in \mathbb{N}$.

Now consider the diagonal subsequence $\{f_k^{(k)}\}_{k\in\mathbb{N}}$. Given $\varepsilon > 0$, let N be large enough that $\frac{1}{N} < \varepsilon$. If $j \geq k > N$, then $f_j^{(j)}$ is one element of the sequence $\{f_n^{(k)}\}_{n\in\mathbb{N}}$, say $f_j^{(j)} = f_n^{(k)}$. Hence

$$\mathrm{d}\big(f_j^{(j)}, f_k^{(k)}\big) \;=\; \mathrm{d}\big(f_n^{(k)}, f_k^{(k)}\big) \;<\; \frac{1}{k} \;<\; \varepsilon.$$

Thus $\{f_k^{(k)}\}_{n\in\mathbb{N}}$ is a Cauchy subsequence of the original sequence. Since E is complete, this subsequence must therefore converge to some element of E.

(b) $\Rightarrow$ (a). Assume that E is sequentially compact. Then by the implication (b) $\Rightarrow$ (c) proved above, we also know that E is complete and totally bounded.

Choose any open cover $\{U_\alpha\}_{\alpha\in J}$ of E. We must show that it contains a finite subcover. The key ingredient in proving this is the following claim.

Claim. There exists a number $\delta > 0$ such that if B is any open ball of radius δ that intersects E, then there is an $\alpha \in J$ such that $B \subseteq U_\alpha$.

To prove the claim, suppose that $\{U_\alpha\}_{\alpha\in J}$ was an open cover of E such that no δ with the required property existed. Then for each $n \in \mathbb{N}$, we could find a open ball B_n with radius $\frac{1}{n}$ that intersects E but is not contained in any U_α. Choose any $f_n \in B_n \cap E$. Since E is sequentially compact, there must be a subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ that converges to an element of E, say $f_{n_k} \to f \in E$. Since $\{U_\alpha\}_{\alpha\in J}$ is a cover of E, we must have $f \in U_\alpha$ for some $\alpha \in J$, and since U_α is open there must exist some $r > 0$ such that $B_r(f) \subseteq U_\alpha$. Now choose k large enough that we have both

$$\frac{1}{n_k} < \frac{r}{3} \qquad \text{and} \qquad \mathrm{d}\big(f, f_{n_k}\big) < \frac{r}{3}.$$

Then it follows that $B_{n_k} \subseteq B_r(f) \subseteq U_\alpha$, which is a contradiction. Hence the claim holds.

To finish the proof, we use the fact that E is totally bounded, and therefore can be covered by finitely many open balls of radius δ. By the claim, each of these balls is contained in some U_α, so E is covered by finitely many of the U_α. $\square$

We will state only a few of the many special properties possessed by functions on compact sets.

Exercise A.70. Let X and Y be topological spaces. If $f\colon X \to Y$ is continuous and $K \subseteq X$ is compact, then $f(K)$ is a compact subset of Y.

Definition A.71. Let (X, d_X) and (Y, d_Y) be metric spaces. We say that a function $f\colon X \to Y$ is *uniformly continuous* if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\forall\, x, y \in X, \quad \mathrm{d}_X(x,y) < \delta \implies \mathrm{d}_Y(f(x),\, f(y)) < \varepsilon.$$

Exercise A.72. Let (X, d_X) and (Y, d_Y) be metric spaces, and let $K \subseteq X$ be compact. Then any continuous function $f\colon K \to Y$ is uniformly continuous.

Additional Problems

A.22. Let X be a topological space.

(a) Prove that every closed subset of a compact subset of X is compact.

(b) Show that if the topology on X is Hausdorff, then every compact subset of X is closed.

A.23. Show that if K is a compact subset of a topological space X, then every infinite subset of K has an accumulation point.

A.24. Show that if E is a totally bounded subset of a metric space X, then its closure $\overline{E}$ is compact. (A set with compact closure is said to be *precompact*.)

A.25. Show that a function $f\colon \mathbb{R}^d \to \mathbb{C}$ is uniformly continuous if and only if $\lim_{a\to 0} \|f - T_a\|_\infty = 0$, where $T_a f(x) = f(x-a)$ is the translation operator.

A.26. (a) Show that every compact subset of a normed linear space X is both closed and bounded.

(b) Show that every closed and bounded subset of $\mathbb{R}^d$ or $\mathbb{C}^d$ is compact.

(c) Use the fact that all norms on a finite-dimensional vector space are equivalent to extend part (b) to arbitrary finite-dimensional vector spaces.

(d) Let H be an infinite-dimensional inner product space. Show that the closed unit ball $\{x \in H : \|x\| \le 1\}$ is closed and bounded but not compact.

A.27. This problem will extend Problem A.26(d) to an arbitrary normed linear space X.

(a) Prove *F. Riesz's Lemma*: If M is a proper, closed subspace of X and $\varepsilon > 0$, then there exists $g \in X$ with $\|g\| = 1$ such that

$$\operatorname{dist}(g, M) \;=\; \inf_{f\in M} \|g - f\| \;>\; 1 - \varepsilon.$$

(b) Prove that the closed unit ball $\{x \in X : \|x\| \le 1\}$ in X is compact if and only if X is finite-dimensional.

A.10 Complete Sequences and a First Look at Schauder Bases

In this section we define complete sequences of vectors in normed spaces. In finite dimensions, these are simply spanning sets. However, in infinite dimensions there are subtle, but important, distinctions between spanning sets, complete sets, and bases. For more details on bases in Banach spaces, we refer to the texts by Singer [Sin70], Lindenstrauss and Tzafriri [LT77], or the introductory text [Hei10].

A.10.1 Span and Closed Span

Definition A.73 (Span). Let A be a subset of a vector space X. The *finite linear span* of A, denoted $\mathrm{span}(A)$, is the set of all finite linear combinations of elements of A:

$$\mathrm{span}(A) \;=\; \biggl\{\sum_{k=1}^{n} \alpha_k f_k : n > 0,\; f_k \in A,\; \alpha_k \in \mathbb{C}\biggr\}.$$

We also refer to the finite linear span of A as the *finite span*, the *linear span*, or simply the *span* of A.

In particular, if A is a countable sequence, say $A = \{f_k\}_{k\in\mathbb{N}}$, then

$$\mathrm{span}\bigl(\{f_k\}_{k\in\mathbb{N}}\bigr) \;=\; \biggl\{\sum_{k=1}^{n} \alpha_k f_k : n > 0,\; \alpha_k \in \mathbb{C}\biggr\}.$$

We will use the following vectors to illustrate many of the concepts in this section.

Definition A.74 (Standard Basis Vectors). Let

$$e_n \;=\; (\delta_{mn})_{m\in\mathbb{N}} \;=\; (0,\dots,0,1,0,0,\dots)$$

denote the sequence that has a 1 in the nth component and zeros elsewhere. We call e_n the nth *standard basis vector*, and refer to $\mathcal{E} = \{e_n\}_{n\in\mathbb{N}}$ as the *standard basis* for ℓ^p (p finite) or c_0 ($p = \infty$).

Of course, this definition begs the question of in what sense the standard basis vectors form a basis. This is answered in Exercise A.81, but for now let us consider the finite span of the standard basis.

Example A.75. The finite span of $\{e_n\}_{n\in\mathbb{N}}$ is $\mathrm{span}\bigl(\{e_n\}_{n\in\mathbb{N}}\bigr) = c_{00}$.

In particular, the finite span of $\{e_n\}_{n\in\mathbb{N}}$ is not ℓ^p for any p and likewise is not c_0. The standard basis vectors do not form a basis for ℓ^p or c_0 in the vector space sense. Instead, in the strict vector space sense of spanning and being finitely independent, the standard basis vectors form a vector space basis for c_{00}.

In a generic vector space, we can form finite linear combinations of vectors, but unless we have a notion of convergence, we cannot take limits or form infinite sums. However, once we impose some extra structure, such as the existence of a norm, these concepts make sense.

Definition A.76 (Closed Span and Complete Sets). Let A be a subset of a normed linear space X.

(a) The *closed finite span* of A, denoted $\overline{\text{span}}(A)$, is the closure of the set of all finite linear combinations of elements of A:

$$\overline{\text{span}}(A) \;=\; \overline{\text{span}(A)} \;=\; \{z \in X : \exists\, y_n \in \text{span}(A) \text{ such that } y_n \to z\}.$$

We also refer to the closed finite span of A as the *closed linear span* or the *closed span* of A.

(b) We say that A is *complete* in X if $\text{span}(A)$ is dense in X, that is, if $\overline{\text{span}}(A) = X$.

There are many other terminologies in use for complete sets, for example, they are also called *total* or *fundamental*.

Example A.77. Let $\mathcal{E} = \{e_n\}_{n\in\mathbb{N}}$ be the standard basis. Taking the closure with respect to the norm $\|\cdot\|_p$, we have $\overline{\text{span}}(\mathcal{E}) = \ell^p$ if $1 \le p < \infty$, and $\overline{\text{span}}(\mathcal{E}) = c_0$ if $p = \infty$.

Beware: The definition of the closed span does NOT imply that

$$\overline{\text{span}}(A) \;=\; \left\{\sum_{k=1}^{\infty} \alpha_k f_k : f_k \in A, \alpha_k \in \mathbb{C}\right\} \qquad \leftarrow \text{This need not hold!}$$

In particular it is NOT true that an arbitrary element of $\overline{\text{span}}(A)$ can be written $f = \sum_{k=1}^{\infty} \alpha_k f_k$ for some $f_k \in A$, $\alpha_k \in \mathbb{C}$ (see Exercise A.83). Instead, to illustrate the meaning of the closed span, consider the case of a countable set $A = \{f_k\}_{k\in\mathbb{N}}$. Here we have

$$\overline{\text{span}}\big(\{f_k\}_{k\in\mathbb{N}}\big) = \left\{f \in X : \exists\, \alpha_{k,n} \in \mathbb{C} \text{ such that } \sum_{k=1}^{n} \alpha_{k,n} f_k \to f \text{ as } n \to \infty\right\}.$$

That is, an element f lies in the closed span of $\{f_k\}_{k\in\mathbb{N}}$ if there exist $\alpha_{k,n} \in \mathbb{C}$ such that

$$\sum_{k=1}^{n} \alpha_{k,n} f_k \to f \quad \text{ as } n \to \infty.$$

In contrast, to say that $f = \sum_{k=1}^{\infty} \alpha_k f_k$ means that

$$\sum_{k=1}^{n} \alpha_k f_k \to f \quad \text{ as } n \to \infty. \tag{A.5}$$

In particular, in order for equation (A.5) to hold, the scalars α_k must be *independent* of n.

A.10.2 Hamel Bases

We have already defined the finite span of a set, and now we recall the definition of finite linear independence.

Definition A.78 (Finite Independence). A subset A of a vector space X is *finitely linearly independent* if every finite subset of A is linearly independent. That is, if $f_1, \dots, f_n$ are any choice of distinct vectors from A, then we must have that

$$\sum_{k=1}^{n} \alpha_k f_k \;=\; 0 \qquad \Longrightarrow \qquad \alpha_1 = \cdots = \alpha_n = 0.$$

We often simply say that A is *linearly independent* or just *independent* to mean that it is finitely linearly independent.

A Hamel basis is an ordinary vector space basis.

Definition A.79 (Hamel Basis). A subset A of a vector space X is a *Hamel basis* or a *vector space basis* for X if it spans and is finitely linearly independent. That is, A is a Hamel basis if $\operatorname{span}(A) = X$ and every finite subset of A is linearly independent.

For example, the standard basis $\mathcal{E} = \{e_n\}_{n\in\mathbb{N}}$ is a Hamel basis for c_{00}, and the set of monomials $\{x^n : n \ge 0\}$ is a Hamel basis for the set $\mathcal{P}$ of all polynomials.

It is shown in Theorem G.3 that the Axiom of Choice implies that every vector space has a basis (in fact, this statement is equivalent to the Axiom of Choice). However, if X is an infinite-dimensional Banach space, then any Hamel basis for X must be uncountable (Exercise C.99). Typically, we cannot explicitly exhibit a Hamel basis for X, i.e., we usually only know that one exists because of the Axiom of Choice. Consequently, Hamel bases are usually not very useful in the setting of normed spaces, and they make only rare appearances in this volume. One interesting consequence of the existence of Hamel bases is that there exist unbounded linear functionals on any infinite-dimensional normed linear space (Problem C.8).

A.10.3 Introduction to Schauder Bases

In a normed space, Hamel bases are unnecessarily restrictive, because they require the *finite* linear span to be the entire space. In contrast, Schauder bases allow "infinite linear combinations," and as such are much more useful in normed spaces. We will be careful to avoid using the term "basis" without qualification, because in the setting of an abstract vector space it usually means a *Hamel basis*, while in the setting of a Banach space it usually means a *Schauder basis*.

Definition A.80 (Schauder Basis). A sequence $\mathcal{F} = \{f_k\}_{k\in\mathbb{N}}$ in a Banach space X is a *Schauder basis* for X if we can write every $f \in X$ as

$$f \;=\; \sum_{k=1}^{\infty} \alpha_k(f)\, f_k \tag{A.6}$$

for a unique choice of scalars $\alpha_k(f)$, where the series converges in the norm of X.

Note that the uniqueness requirement implies that $f_k \neq 0$ for every k.

Exercise A.81. Show that the standard basis $\mathcal{E} = \{e_n\}_{n\in\mathbb{N}}$ is a Schauder basis for ℓ^p for each $1 \le p < \infty$, and $\mathcal{E}$ is a Schauder basis for c_0 with respect to ℓ^∞-norm.

Every Schauder basis is both complete and finitely linearly independent. However, the next exercise shows that *the converse fails in general.* For this example we will need the following useful theorem on approximation by polynomials (see Theorem 1.89 for proof). Given $a < b$, we set

$$C[a,b] \;=\; \{f\colon [a,b] \to \mathbb{C} \;:\; f \text{ is continuous}\}. \tag{A.7}$$

This is a Banach space under the uniform norm.

Theorem A.82 (Weierstrass Approximation Theorem). *If $f \in C[a,b]$ and $\varepsilon > 0$, then there exists a polynomial p such that*

$$\|f - p\|_\infty \;=\; \sup_{x\in[a,b]} |f(x) - p(x)| \;<\; \varepsilon.$$

Exercise A.83. Use the Weierstrass Approximation Theorem to show that the set of monomials $\{x^k\}_{k=0}^\infty$ is complete and finitely linearly independent in $C[a,b]$. Prove that there exist functions $f \in C[a,b]$ that cannot be written as $f(x) = \sum_{k=0}^\infty \alpha_k x^k$ with convergence of the series in the uniform norm. Consequently $\{x^k\}_{k=0}^\infty$ is not a Schauder basis for $C[a,b]$.

A complete discussion of Schauder bases requires the Hahn–Banach and Uniform Boundedness Theorems, and therefore we will postpone additional discussion of the relations and distinctions between bases and complete linearly independent sets until Section C.15. One interesting fact that we will see there is that the uniqueness requirement in the definition of a Schauder basis implies that the functionals α_k appearing in equation (A.6) must be continuous, and hence belong to the dual space of X.

We end this section with one important exercise.

Exercise A.84. Let X be a Banach space. Show that if there exists a countable subset $\{f_n\}_{n\in\mathbb{N}}$ in X that is complete, then X is separable.

In particular, a Banach space that has a Schauder basis must be separable. The question of whether every separable Banach space has a basis was a longstanding open problem known as the *Basis Problem.* It was finally shown by Enflo [Enf73] that there exist separable reflexive Banach spaces that have no Schauder bases!

A.11 Unconditional Convergence

Recall that a series $\sum_{n=1}^{\infty} f_n$ in a normed space X converges and equals a vector f if the sequence of partial sums $\sum_{n=1}^{N} f_n$ converges to f as $N \to \infty$. However, the ordering of the f_n can be important, for if we rearrange the series, it may no longer converge. A series that converges regardless of ordering possesses important stability properties that we will review in this section.

Definition A.85 (Unconditionally Convergent Series). Let X be a normed linear space and let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of elements of X. Then the series $\sum_{n=1}^{\infty} f_n$ is said to *converge unconditionally* if every rearrangement of the series converges, i.e., if for each bijection $\sigma\colon \mathbb{N} \to \mathbb{N}$ the series

$$\sum_{n=1}^{\infty} f_{\sigma(n)}$$

converges in X.

Thus, a series $\sum_{n=1}^{\infty} f_n$ converges unconditionally if and only if for each bijection $\sigma\colon \mathbb{N} \to \mathbb{N}$, the partial sums $\sum_{n=1}^{N} f_{\sigma(n)}$ converge to some vector as $N \to \infty$. Note that while we do not explicitly require in this definition that the partial sums converge to the *same* limit for each bijection σ, Exercise A.90 shows that this must in fact happen.

The method of the next example can be used to show that if X is a *finite-dimensional* normed space, then absolute convergence is equivalent to unconditional convergence. However, this equivalence fails in any infinite dimensional normed space. That is, if X is infinite-dimensional, then there will exist series that converge unconditionally but not absolutely.

Example A.86. To illustrate the importance of unconditional convergence, consider the Banach space $X = \mathbb{C}$. The *harmonic series* $\sum_{n=1}^{\infty} \frac{1}{n}$ does not converge, as its partial sums are unbounded. On the other hand, the *alternating harmonic series* $\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n}$ does converge (in fact, its partial sums converge to $\ln 2$). While this latter series converges, it does not converge absolutely.

Now consider what happens if we change the order of summation. For $n \in \mathbb{N}$, let $p_n = \frac{1}{2n-1}$ and $q_n = \frac{1}{2n}$, i.e., the p_n are the positive terms from the alternating series and the q_n are the absolute values of the negative terms. Each series $\sum p_n$ and $\sum q_n$ diverges to infinity. Hence there must exist an $m_1 > 0$ such that

$$p_1 + \cdots + p_{m_1} \ > \ 1.$$

Also, there must exist an $m_2 > m_1$ such that

$$p_1 + \cdots + p_{m_1} - q_1 + p_{m_1+1} + \cdots + p_{m_2} \ > \ 2.$$

Continuing in this way, we can create a rearrangement

$$p_1 + \cdots + p_{m_1} - q_1 + p_{m_1+1} + \cdots + p_{m_2} - q_2 + \cdots$$

of $\sum (-1)^n \frac{1}{n}$ that diverges to $+\infty$. Likewise, we can construct a rearrangement that diverges to $-\infty$, converges to any given real number r, or simply oscillates without ever converging.

A modification of the argument above gives us the following theorem for series of scalars.

Theorem A.87. *A series of scalars converges absolutely if and only if it converges unconditionally. That is, given $c_n \in \mathbb{C}$,*

$$\sum_{n=1}^{\infty} c_n \text{ converges unconditionally} \quad \Longleftrightarrow \quad \sum_{n=1}^{\infty} |c_n| < \infty.$$

One direction of Theorem A.87 generalizes to all Banach spaces.

Exercise A.88. Let X be a Banach space. Show that if a series $\sum_{n=1}^{\infty} f_n$ converges absolutely in X then it converges unconditionally.

The converse of Exercise A.88 fails in general. For example, if $\{e_n\}_{n\in\mathbb{N}}$ is any orthonormal sequence in a Hilbert space H, then the series $\sum \frac{1}{n} e_n$ converges unconditionally but not absolutely. It can be shown that absolute and unconditional convergence are equivalent *only* for finite-dimensional spaces (this is the *Dvoretzky–Rogers Theorem*, see [LT77]).

The following theorem, which we state without proof, gives several reformulations of unconditional convergence. For this result, recall the definition of directed set given in Definition A.42, and note that the set I consisting of all finite subsets of $\mathbb{N}$ forms a directed set with respect to inclusion of sets.

Theorem A.89. *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in a Banach space X. Then the following statements are equivalent.*

(a) *$\sum_{n=1}^{\infty} f_n$ converges unconditionally.*

(b) *The net $\bigl\{\sum_{n\in F} f_n : F \subseteq \mathbb{N}, F \text{ finite}\bigr\}$ converges in X. That is, there exists an $f \in X$ such that for every $\varepsilon > 0$, there is a finite $F_0 \subseteq \mathbb{N}$ such that*

$$F_0 \subseteq F \subseteq \mathbb{N},\ F \text{ finite} \implies \Bigl\| f - \sum_{n\in F} f_n \Bigr\| < \varepsilon.$$

(c) *If $(c_n)_{n\in\mathbb{N}}$ is any bounded sequence of scalars, then $\sum_{n=1}^{\infty} c_n f_n$ converges in X.*

Exercise A.90. Use Theorem A.89 to show that if a series $\sum_{n=1}^{\infty} f_n$ converges unconditionally in a Banach space X, then there exists a single $f \in X$ such that for every bijection $\sigma\colon \mathbb{N} \to \mathbb{N}$ we have $\sum_{n=1}^{\infty} f_{\sigma(n)} = f$.

A.12 Orthogonality

In this section we review some of the special properties of inner product spaces, especially in relation to orthogonality.

A.12.1 Orthogonality and the Pythagorean Theorem

Definition A.91. Let H be an inner product space, and let I be an arbitrary index set.

(a) Vectors f, $g \in H$ are *orthogonal*, denoted $f \perp g$, if $\langle f, g\rangle = 0$.

(b) A collection of vectors $\{f_i\}_{i\in I}$ is *orthogonal* if $\langle f_i, f_j\rangle = 0$ whenever $i \neq j$.

(c) A collection of vectors $\{f_i\}_{i\in I}$ is *orthonormal* if it is orthogonal and each vector is a unit vector, i.e., $\langle f_i, f_j\rangle = \delta_{ij}$.

For example, the standard basis is an orthonormal sequence in ℓ^2.

Exercise A.92 (Pythagorean Theorem). Show that if $f_1, \dots, f_n \in H$ are orthogonal, then

$$\left\|\sum_{k=1}^{n} f_k\right\|^2 = \sum_{k=1}^{n} \|f_k\|^2.$$

The existence of a notion of orthogonality gives inner product spaces a much "simpler" structure than general Banach spaces. We will derive some of the basic properties of inner product spaces below, most of which are straightforward consequences of orthogonality. Some (but not all) of these results have analogues for general Banach spaces. However, even for the results that do have analogues, the corresponding proofs in the non-Hilbert space setting are usually much more complicated or are nonconstructive (see the discussion of the Hahn–Banach Theorem in Section C.10).

A.12.2 Orthogonal Direct Sums

Definition A.93 (Orthogonal Direct Sum). Let M, N be closed subspaces of a Hilbert space H.

(a) The *direct sum* of M and N is $M + N = \{f + g : f \in M, g \in N\}$.

(b) We say that M and N are *orthogonal subspaces*, denoted $M \perp N$, if $f \perp g$ for every $f \in M$ and $g \in N$.

(c) If M, N are orthogonal subspaces in H, then we call their direct sum the *orthogonal direct sum* of M and N, and denote it by $M \oplus N$.

Exercise A.94. Show that if M, N are closed, orthogonal subspaces of H, then $M \oplus N$ is a closed subspace of H.

A.12.3 Orthogonal Projections and Orthogonal Complements

Definition A.95. Let X be a normed linear space and fix $A \subseteq H$. The *distance* from a point $f \in H$ to the set A is

$$\operatorname{dist}(f, A) \;=\; \inf\bigl\{\|f-g\| : g \in A\bigr\}.$$

Given a closed convex subset K of a Hilbert space H and given any $x \in H$, there is a unique point in K that is closest to x.

Exercise A.96 (Closest Point Property). If H is a Hilbert space and K is a nonempty closed, convex subset of H, then given any $h \in H$ there exists a unique point $k_0 \in K$ that is closest to h. That is, there is a unique point $k_0 \in K$ such that

$$\|h-k_0\| \;=\; \operatorname{dist}(h, K) \;=\; \inf\bigl\{\|h-k\| : k \in K\bigr\}.$$

Applying the closest point theorem to the particular case of closed subspaces gives us the existence of *orthogonal projections* in a Hilbert space.

Definition A.97 (Orthogonal Projection). Let M be a closed subspace of a Hilbert space H. For each $h \in H$, the unique point $p \in M$ closest to h is called the *orthogonal projection* of h onto M.

Orthogonal projections can be characterized as follows.

Exercise A.98. Let M be a closed subspace of a Hilbert space H. Given $h \in H$, show that the following statements are equivalent.

(a) $h = p + e$ where p is the (unique) point in M closest to h.

(b) $h = p + e$ where $p \in M$ and $e \perp M$ (i.e., $e \perp f$ for every $f \in M$).

Definition A.99 (Orthogonal Complement). Let A be a subset (not necessarily a subspace) of a Hilbert space H. The *orthogonal complement* of A is

$$A^\perp \;=\; \bigl\{f \in H : f \perp A\bigr\} \;=\; \bigl\{f \in H : \langle f, g\rangle = 0 \text{ for all } g \in A\bigr\}.$$

In the terminology of orthogonal complements, a vector p is the orthogonal projection of f onto a closed subspace M if and only if $f = p + e$ where $p \in M$ and $e \in M^\perp$.

Exercise A.100. Show that if A is an arbitrary subset of a Hilbert space H, then $A^\perp$ is a closed subspace of H (even if A is not), and $(A^\perp)^\perp = \overline{\operatorname{span}}(A)$.

Hence, if M is a closed subspace of a Hilbert space H then $(M^\perp)^\perp = M$. Further, $H = M \oplus M^\perp$, the orthogonal direct sum of M and $M^\perp$.

The next exercise will allow us to give a useful equivalent formulation of completeness of a sequence in a Hilbert space.

Exercise A.101. Let A be a subset of a Hilbert space H. Show that A is complete if and only if $A^\perp = \{0\}$, i.e., the only vector orthogonal to every element of A is the zero vector.

A.13 Orthogonality and Complete Sequences

In a Hilbert space, the combination of completeness and orthonormality of a sequence $\{e_n\}_{n\in\mathbb{N}}$ leads to especially nice series representations of vectors in H in terms of the vectors e_n. We will explore this topic in this section.

The following exercise summarizes some basic results connected to convergence of series of orthonormal vectors.

Exercise A.102. Let $\{e_n\}_{n\in\mathbb{N}}$ be any orthonormal sequence in a Hilbert space H. Show that the following statements hold.

(a) Bessel's Inequality: $\sum_{n=1}^{\infty} |\langle f, e_n\rangle|^2 \le \|f\|^2$.

(b) If $f = \sum_{n=1}^{\infty} c_n e_n$ converges, then $c_n = \langle f, e_n\rangle$.

(c) $\sum_{n=1}^{\infty} c_n e_n$ converges $\iff$ $\sum_{n=1}^{\infty} |c_n|^2 < \infty$.

(d) If $\sum_{n=1}^{\infty} c_n e_n$ converges then it converges *unconditionally*, i.e., it converges regardless of the ordering of the indices (see Section A.11).

(e) $f \in \overline{\operatorname{span}}\bigl(\{e_n\}_{n\in\mathbb{N}}\bigr) \iff f = \sum_{n=1}^{\infty} \langle f, e_n\rangle\, e_n$.

(f) If $f \in H$, then $p = \sum_{n=1}^{\infty} \langle f, e_n\rangle\, e_n$ is the orthogonal projection of f onto $\overline{\operatorname{span}}\bigl(\{e_n\}_{n\in\mathbb{N}}\bigr)$.

Now we characterize those sequences in a Hilbert space that are both complete and orthonormal.

Exercise A.103. Let $\{e_n\}_{n\in\mathbb{N}}$ be any orthonormal sequence in a Hilbert space H. Show that the following statements are equivalent.

(a) $\{e_n\}_{n\in\mathbb{N}}$ is complete.

(b) For each $f \in H$ there exist unique scalars $(c_n)_{n\in\mathbb{N}}$ such that $f = \sum_{n=1}^{\infty} c_n e_n$.

(c) For each $f \in H$, $f = \sum_{n=1}^{\infty} \langle f, e_n\rangle\, e_n$.

(d) Plancherel's Equality: For each $f \in H$, $\|f\|^2 = \sum_{n=1}^{\infty} |\langle f, e_n\rangle|^2$.

(e) Parseval's Equality: For each $f, g \in H$, $\langle f, g\rangle = \sum_{n=1}^{\infty} \langle f, e_n\rangle\, \langle e_n, g\rangle$.

Thus, *for a countable orthonormal sequence*, completeness implies the existence of series expansions of vectors. This need not be true for arbitrary complete sequences, even in a Hilbert space (Problem A.28).

Definition A.104 (Orthonormal Basis). Let H be a Hilbert space. An orthonormal sequence $\{e_n\}_{n\in\mathbb{N}}$ which satisfies the equivalent conditions of Exercise A.103 is called an *orthonormal basis* for H.

Remark A.105. Recall from Definition A.80 that a sequence $\{f_n\}_{n\in\mathbb{N}}$ in a Banach space X is called a *Schauder basis* if every vector $f \in X$ can be written $f = \sum_{n=1}^{\infty} c_n(f) f_n$ for a unique choice of scalars $c_n(f)$. In this terminology, an *orthonormal* sequence is complete if and only if it is a Schauder basis. However, it is important to emphasize that an arbitrary complete sequence *need not* be a Schauder basis. Thus, in a Hilbert space we have:

$$\text{complete orthonormal sequence} \quad \begin{array}{c}\Longrightarrow \\ \not\Longleftarrow\end{array} \quad \text{Schauder basis.}$$

We have been considering *countable* orthonormal bases. By Exercise A.84, any Hilbert space that has a countable orthonormal basis must be separable. Conversely, every separable Hilbert space has an orthonormal basis (see Problem A.31). We show explicitly in Theorem 2.22 that $\{e^{2\pi inx}\}_{n\in\mathbb{Z}}$ forms an orthonormal basis for $L^2[0,1)$.

The results of this section can be extended to uncountable orthonormal sequences in nonseparable Hilbert spaces (such as $\ell^2(I)$ when I is uncountable), but then we must be extremely careful regarding the use of the terminology "basis," as in much of the Banach space literature the terminology "basis" is reserved for a *countable* sequence that forms a Schauder basis. We summarize without proof the results that hold for uncountable orthonormal sequences in Hilbert spaces (see [Con90] for details).

Theorem A.106. *Let H be a Hilbert space and let I be an index set. If $\{e_i\}_{i\in I}$ is an orthonormal set in H, then the following statements hold.*

(a) *If $f \in H$ then $\langle f, e_i\rangle \ne 0$ for at most countably many $i \in I$.*

(b) *For each $f \in H$, $\sum_{i\in I} |\langle f, e_i\rangle|^2 \le \|f\|^2$.*

(c) *For each $f \in H$, the series $p = \sum_{i\in I} \langle f, e_i\rangle\, e_i$ converges with respect to the net of finite subsets of I, and p is the orthogonal projection of f onto $\overline{\mathrm{span}}\big(\{e_i\}_{i\in I}\big)$.*

Theorem A.107. *Let $\{e_i\}_{i\in I}$ be an orthonormal set in a Hilbert space H. Then the following statements are equivalent.*

(a) *$\{e_i\}_{i\in I}$ is complete.*

(b) *For each $f \in H$ we have $f = \sum_{i\in I} \langle f, e_i\rangle\, e_i$, where the series converges with respect to the net of finite subsets of I.*

(c) *For each $f \in H$, $\|f\|^2 = \sum_{i\in I} |\langle f, e_i\rangle|^2$.*

Note that, because of orthogonality, the series that appear in both Theorems A.106 and A.107 contain only countably many nonzero terms. If we restrict to just those countably many nonzero terms, then, in the language of Section A.11, the series converge unconditionally.

Additional Problems

A.28. We say that a sequence $\{f_n\}_{n\in\mathbb{N}}$ in a Banach space X is ω-*dependent* if there exist scalars c_n, not all zero, such that $\sum_{n=1}^{\infty} c_n f_n = 0$, where the series converges in the norm of X. A sequence is ω-*independent* if it is not ω-dependent.

(a) Show that every Schauder basis is complete and ω-independent.

(b) Let $\alpha, \beta \in \mathbb{C}$ be fixed nonzero scalars such that $|\alpha| > |\beta|$. Let $\{e_n\}_{n\in\mathbb{N}}$ be the standard basis for ℓ^2, and define $x_0 = e_1$ and $x_n = \alpha e_n + \beta e_{n+1}$ for $n \in \mathbb{N}$. Prove that $\{x_k\}_{k\geq 0}$ is complete and finitely independent in ℓ^2, but is not ω-independent and therefore is not a Schauder basis for ℓ^2.

A.29. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in a Hilbert space H. Prove that the following two statements are equivalent.

(a) For each $m \in \mathbb{N}$ we have $f_m \notin \overline{\operatorname{span}}\{f_n\}_{n\neq m}$ (such a sequence is said to be *minimal*).

(b) There exists a sequence $\{g_n\}_{n\in\mathbb{N}}$ in H such that $\langle f_m, g_n\rangle = \delta_{mn}$ for all m, $n \in \mathbb{N}$ (we say that sequences $\{f_n\}_{n\in\mathbb{N}}$ and $\{g_n\}_{n\in\mathbb{N}}$ satisfying this condition are *biorthogonal*).

Show further that in case statements (a), (b) hold, the sequence $\{g_n\}_{n\in\mathbb{N}}$ is unique if and only if $\{f_n\}_{n\in\mathbb{N}}$ is complete.

Remark: Exercise C.111 establishes an analogous result for Banach spaces.

A.30. Formulate the Gram–Schmidt orthogonalization procedure in an arbitrary inner product space, and use it to show that any infinite-dimensional inner product space contains an infinite orthonormal sequence $\{e_n\}_{n\in\mathbb{N}}$.

A.31. Use the Axiom of Choice in the form of Zorn's Lemma (Axiom G.2) to show that every Hilbert space has a complete orthonormal subset. In particular, every separable Hilbert space has an orthonormal basis.

A.14 Urysohn's Lemma

Urysohn's Lemma is a general result that holds in a large class of topological spaces (specifically, the *normal* topological spaces, which include all metric spaces). It states that if A and B are disjoint closed subsets of a normal topological space X, then there exists a continuous function $f\colon X \to [0,1]$ that is identically 0 on A and identically 1 on B. We prove here a version of Urysohn's Lemma for $\mathbb{R}^d$ (and the same simple proof can be used in any metric space). A more refined version of Urysohn's Lemma for real line is proved in Chapter 1 (see Theorem 1.60).

First, we need the following lemma.

Lemma A.108. *If $E \subseteq \mathbb{R}^d$ is nonempty, then*

$$f(x) = \operatorname{dist}(x, E) = \inf\bigl\{|x - z| : z \in E\bigr\}$$

is uniformly continuous on $\mathbb{R}^d$.

Proof. Fix $\varepsilon > 0$. Choose any $x, y \in \mathbb{R}^d$ with $|x-y| < \varepsilon/2$. By definition, there exist $a, b \in E$ such that $|x-a| < \operatorname{dist}(x,E)+\varepsilon/2$ and $|y-b| < \operatorname{dist}(y,E)+\varepsilon/2$. Hence

$$\begin{aligned} f(y) = \operatorname{dist}(y, A) &\le |y - a| \\ &\le |y - x| + |x - a| \\ &< \frac{\varepsilon}{2} + \operatorname{dist}(x, E) + \frac{\varepsilon}{2} \\ &= f(x) + \varepsilon. \end{aligned}$$

Similarly $f(x) < f(y) + \varepsilon$, so $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \varepsilon/2$. □

Theorem A.109 (Urysohn's Lemma). *If E, F are disjoint closed subsets of $\mathbb{R}^d$, then there exists a continuous function $\theta \colon \mathbb{R}^d \to \mathbb{R}^d$ such that*

(a) $0 \le \theta \le 1$,

(b) $\theta = 0$ *on E, and*

(c) $\theta = 1$ *on F.*

Proof. Because E is closed, if $x \notin E$ then we have $\operatorname{dist}(x, E) > 0$. Also, by Lemma A.108, $\operatorname{dist}(x, E)$ and $\operatorname{dist}(x, F)$ are each continuous functions of x. It follows that the function

$$\theta(x) = \frac{\operatorname{dist}(x, E)}{\operatorname{dist}(x, E) + \operatorname{dist}(x, F)}$$

has the required properties. □