245B Real Analysis: Some solutions

ho boon suan

January 2, 2022 to April 25, 2022, 17:00

In this document, I produce some solutions to exercises as I work through Terence Tao's UCLA sequence on graduate real analysis. The lecture notes for the second and third parts of his three part sequence 245ABC are collected in his book *An Epsilon of Room, I: Real Analysis*.

I have attempted most exercises, but for some where I got stuck for too long, I looked up solutions online. I have indicated my references in square brackets at the beginning of such solutions.

Contents

Oh! If only someone
would give me time, time,
time to do everything properly,
to read everything at my own tempo,
to take it apart and put it together again.

— Karl Barth (1922)

1.1. A quick review of measure and integration theory

The ultimate measure of a man is not where
he stands in moments of comfort and convenience,
but where he stands at times of challenge and controversy.

— Martin Luther King, Jr., *Strength to Love* (1963)

Exercise 1.1.1. We use a kind of 'structural induction' to prove the claim (see 245A, Remark 1.4.15). We recall the principle here for convenience.

I wonder if it would be sleeker to do this via transfinite induction. I haven't learned the details of this method yet though, so I won't try it for now.

Remark. If $\mathcal{F}$ is a family of sets in X, and $P(E)$ is a property of sets $E \subset X$ which obeys the following axioms:

(i) $P(\varnothing)$ is true.

(ii) $P(E)$ is true for all $E \in \mathcal{F}$.

(iii) If $P(E)$ is true for some $E \subset X$, then $P(X \setminus E)$ is true also.

(iv) If $E_1, E_2, \ldots \subset X$ are such that $P(E_n)$ is true for all n, then $P(\bigcup_{n=1}^\infty E_n)$ is true also.

Then one can conclude that $P(E)$ is true for all $E \in \langle \mathcal{F} \rangle$. Indeed, the set of all E for which $P(E)$ holds is a σ-algebra containing $\mathcal{F}$.

We now prove that a continuous function f between topological spaces X and Y is Borel measurable, by using the remark above with $\mathcal{F}$ being the family of open sets in Y, and $P(E)$ the property that $f^{-1}(E)$ is Borel measurable in X. Claim (i) holds as $f^{-1}(\varnothing) = \varnothing$. Claim (ii) holds by continuity. Claim (iii) follows from the identity $f^{-1}(Y \setminus E) = X \setminus f^{-1}(E)$. Finally, claim (iv) follows from the fact that $f^{-1}(\bigcup_{n=1}^\infty E_n) = \bigcup_{n=1}^\infty f^{-1}(E_n)$.

Remark. The Borel σ-algebra $\mathcal{B}[S]$ of a subspace $S \subset X$ is equal to the Borel σ-algebra of X restricted to S. That is, $\mathcal{B}[S] = \mathcal{B}[X]\lfloor_S$. Indeed, they are both generated by sets of the form $U \cap S$, where $U \subset X$ is open. (Be careful not to confuse the notations $\mathcal{B}[X]$ and $\mathcal{B}[\mathcal{F}]$. The first refers to the σ-algebra generated by the open sets of a topological space X, and the second is the smallest σ-algebra containing a family of sets $\mathcal{F} \subset \mathcal{P}(X)$.)

Exercise 1.1.2. We wish to prove that $\mathcal{B}[M]$ is maximal such that

$$\mathcal{B}[M]\lfloor_{U_\alpha} = \pi_\alpha^{-1}(\mathcal{B}[\mathbf{R}^n]\lfloor_{V_\alpha})$$

for all α. By exercise 1.1.1, we see that a homeomorphism between topological spaces induce a bijection between their σ-algebras. Thus it suffices to prove maximality. Suppose $\mathcal{X}$ is a σ-algebra on M satisfying the above identities, so that

$$\mathcal{X}\lfloor_{U_\alpha} = \pi_\alpha^{-1}(\mathcal{B}[\mathbf{R}^n]\lfloor_{V_\alpha}) = \mathcal{B}[M]\lfloor_{U_\alpha} .$$

Then, it suffices to show that any element of $\mathcal{X}$ is a countable union of sets, each belonging to some $\mathcal{X}\lfloor_{U_\alpha}$. By the second countability of

M, we may choose U_{α_i} such that their union covers $X \in \mathcal{X}$. Thus $X = \bigcup_i X \cap U_{\alpha_i}$, so that $X \cap U_{\alpha_i} \in \mathcal{B}[M]\lfloor_{U_{\alpha_i}}$, and we are done.

Exercise 1.1.3. Let $\mathcal{X}$ be a σ-algebra on a finite set X. We define a map $X \to \mathcal{X}$ sending $x \in X$ to the intersection of all sets in $\mathcal{X}$ containing x. We prove that the image of this map is a partition of X, and that $\mathcal{X}$ arises from this partition. Clearly the image covers X. Suppose x and y get sent to sets S_x and S_y with non-empty intersection. Then $x \in S_x \cap (X \setminus S_y) \subsetneq S_x$, contradicting the minimality of S_x. Thus the sets form a partition. Given a set $X \in \mathcal{X}$, we see that $X = \bigcup_{x \in X} S_x$, where the sets in the union are either identical or disjoint. Discarding repeated sets, we obtain the claim.

Exercise 1.1.4. Let $(X_\alpha)_{\alpha \in A}$ be an at most countable family of second countable topological spaces. We prove that

$$\mathcal{B}\Big[\prod_{\alpha \in A} X_\alpha\Big] = \prod_{\alpha \in A} \mathcal{B}[X_\alpha].$$

Let $\prod_{\alpha \in A} B_\alpha \in \prod_{\alpha \in A} \mathcal{B}[X_\alpha]$. Since the projections $\pi_\beta \colon \prod_{\alpha \in A} X_\alpha \to X_\beta$ are Borel measurable (by definition of the product σ-algebra), the sets $\pi_\alpha^{-1}(B_\alpha)$ belong to $\mathcal{B}[\prod_{\alpha \in A} X_\alpha]$, and so

$$\prod_{\alpha \in A} B_\alpha = \bigcap_{\alpha \in A} \pi_\alpha^{-1}(B_\alpha) \in \mathcal{B}\Big[\prod_{\alpha \in A} X_\alpha\Big]$$

as needed.

For the forward inclusion $\subset$, we see that since the σ-algebra $\mathcal{B}[\prod_{\alpha \in A} X_\alpha]$ is generated by open sets in $\prod_{\alpha \in A} X_\alpha$, it suffices to prove that these open sets belong to $\prod_{\alpha \in A} \mathcal{B}[X_\alpha]$. Expanding the definition of $\prod_{\alpha \in A} \mathcal{B}[X_\alpha]$, we have

$$\prod_{\alpha \in A} \mathcal{B}[X_\alpha] = \bigvee_{\alpha \in A} \pi_\alpha^{-1}(\mathcal{B}[X_\alpha]) = \mathcal{B}\Big[\bigcup_{\alpha \in A} \pi_\alpha^{-1}(\mathcal{B}[X_\alpha])\Big].$$

For each X_α, we let $\mathcal{B}_\alpha$ be a countable base. Then

$$\mathcal{B} := \left\{ \prod_{\alpha \in A} U_\alpha : \begin{array}{c} U_\alpha = X_\alpha \text{ for all but finitely many } \alpha, \\ \text{and if } U_\alpha \neq X_\alpha, \text{ then } U_\alpha \in \mathcal{B}_\alpha. \end{array} \right\}$$

is a countable base for $\prod_{\alpha \in A} X_\alpha$. It remains to show that $\prod_{\alpha \in A} U_\alpha \in \mathcal{B}$ belongs to $\mathcal{B}[\bigcup_{\alpha \in A} \pi_\alpha^{-1}(\mathcal{B}[X_\alpha])]$. Writing $U_{\alpha_1}, \dots, U_{\alpha_n}$ for the finitely many nontrivial sets in the product $\prod_{\alpha \in A} U_\alpha$, we see that such a set is a finite intersection

$$\prod_{\alpha \in A} U_\alpha = \bigcap_{1 \le i \le n} \pi_{\alpha_i}^{-1}(U_{\alpha_i}) \in \mathcal{B}\Big[\bigcup_{\alpha \in A} \pi_\alpha^{-1}(\mathcal{B}[X_\alpha])\Big],$$

and thus the result follows.

Exercise 1.1.5. We proceed via structural induction (see the remark on page 1). Given $x \in X$, we write $E_x := \{y \in Y : (x, y) \in E\}$, and we call it a *slice* of E (we define E^y similarly). Claim (i) is trivial as all slices of the empty set are empty. For claim (ii), we see that the family $\mathcal{X} \times \mathcal{Y}$ of measurable sets has measurable slices — indeed, given $A \times B \in \mathcal{X} \times \mathcal{Y}$ and $x \in A$, any slice $(A \times B)_x \subset Y$ is either $\varnothing$ or B, and is measurable in both cases. Claim (iii) follows from how

$$((X \times Y) \setminus E)_x = \{y \in Y : (x, y) \notin E\} = Y \setminus E_x.$$

Finally, claim (iv) follows from the fact that

$$\Big(\bigcup_{n=1}^{\infty} E_n\Big)_x = \Big\{y \in Y : (x, y) \in \bigcup_{n=1}^{\infty} E_n\Big\} = \bigcup_{n=1}^{\infty} (E_n)_x.$$

Thus the result holds for $x \in X$; the proof is analogous for $y \in Y$.

Exercise 1.1.6. (i) Countable additivity implies finite additivity by setting $E_n := \varnothing$ for $n \geq N$. Therefore, if $E \subset F$, then $\mu(F) = \mu(E) + \mu(F \setminus E)$, and the result follows from nonnegativity of measure.

(ii) Define $E'_n := E_n \setminus \bigcup_{k=1}^{n-1} E_k$ for $n \geq 1$. The sets E'_n are disjoint with $E_n = \bigcup_{k=1}^{n} E'_k$, and consequently $\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} E'_n$. Thus, by countable additivity and (i), we have

$$\mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) = \mu\Big(\bigcup_{n=1}^{\infty} E'_n\Big) = \sum_{n=1}^{\infty} \mu(E'_n) \leq \sum_{n=1}^{\infty} \mu(E_n).$$

(iii) Since $E_n \subset \bigcup_{k=1}^{\infty} E_k$ for $n \geq 1$, monotonicity implies that $\mu(E_n) \leq \mu(\bigcup_{k=1}^{\infty} E_k)$ for $n \geq 1$, so that $\lim_{n\to\infty} \mu(E_n) \leq \mu(\bigcup_{n=1}^{\infty} E_n)$. Conversely, writing $E'_n := E_n \setminus \bigcup_{k=1}^{n-1} E_k$, we may compute

$$\begin{aligned}
\lim_{n\to\infty} \mu(E_n) &= \lim_{n\to\infty} \mu\Big(\bigcup_{k=1}^{n} E'_k\Big) \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} \mu(E'_k) \\
&= \sum_{n=1}^{\infty} \mu(E'_n) \\
&\geq \mu\Big(\bigcup_{n=1}^{\infty} E'_n\Big) \\
&= \mu\Big(\bigcup_{n=1}^{\infty} E_n\Big).
\end{aligned}$$

(iv) Apply (iii) to the sequence $\varnothing \subset E_1 \setminus E_2 \subset E_1 \setminus E_3 \subset \ldots$ to obtain the identity

$$\begin{aligned}
\mu(E_1) - \mu\Big(\bigcap_{n=1}^{\infty} E_n\Big) &= \mu\Big(\bigcup_{n=1}^{\infty} E_1 \setminus E_n\Big) \\
&= \lim_{n\to\infty} \mu(E_1 \setminus E_n) \\
&= \lim_{n\to\infty} (\mu(E_1) - \mu(E_n)).
\end{aligned}$$

Note that the claim fails if $\mu(E_1) = +\infty$, consider for example $(1, +\infty) \subset (2, +\infty) \subset (3, +\infty) \subset \ldots$, where each set has infinite measure, but the intersection is empty and thus has zero measure.

Exercise 1.1.7. Given a measure space $(X, \mathcal{X}, \mu)$, we define a new σ-algebra $\overline{\mathcal{X}}$ to be the intersection of all σ-algebras containing $\mathcal{X}$ as well as all subsets of null sets. By definition, this new measurable space $(X, \overline{\mathcal{X}}, \mu)$ is the unique minimal complete refinement of $(X, \mathcal{X}, \mu)$. If a set A is equal a.e. to a set $B \in \mathcal{X}$, then their symmetric difference $A \triangle B$ is a sub-null set, and so $A = (A \triangle B) \triangle B \in \overline{\mathcal{X}}$. Conversely, we may use structural induction. For (i), the empty set belongs to all σ-algebras. For (ii), this is true for all sub-null sets and all elements of $\mathcal{X}$. For (iii), if $E = F$ a.e., then $X \setminus E = X \setminus F$ a.e.. For (iv), if $E_n = F_n$ a.e., then $\bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} F_n$ a.e., since the countable union of sub-null sets is sub-null. Thus the result follows.

Exercise 1.1.8. [Halmos, *Measure Theory*, page 56–57, Theorem D] Suppose $E \subset X$ with $\mu(E) < \infty$. By definition of μ, there exist sets $(A_n)_{n=1}^{\infty}$ in $\mathcal{A}$ such that $E \subset \bigcup_{n=1}^{\infty} A_n$ and $\mu(\bigcup_{n=1}^{\infty} A_n) \leq \mu(E) + \epsilon/2$. By monotone convergence, we have $\lim_{N\to\infty} \mu(\bigcup_{n=1}^{N} A_n) = \mu(\bigcup_{n=1}^{\infty} A_n)$. Thus we may choose large N for which $\mu(\bigcup_{n=1}^{\infty} A_n) \leq \mu(\bigcup_{n=1}^{N} A_n) + \epsilon/2$. Since

$$\begin{aligned} \mu\Big(E \setminus \bigcup_{n=1}^{N} A_n\Big) &\leq \mu\Big(\bigcup_{n=1}^{\infty} A_n \setminus \bigcup_{n=1}^{N} A_n\Big) \\ &= \mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) - \mu\Big(\bigcup_{n=1}^{N} A_n\Big) \\ &\leq \epsilon/2 \end{aligned}$$

and

$$\begin{aligned} \mu\Big(\bigcup_{n=1}^{N} A_n \setminus E\Big) &\leq \mu\Big(\bigcup_{n=1}^{\infty} A_n \setminus E\Big) \\ &= \mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) - \mu(E) \\ &\leq \epsilon/2, \end{aligned}$$

we conclude that

$$\mu\Big(E \triangle \bigcup_{n=1}^{N} A_n\Big) \leq \epsilon$$

as desired. [To do: complete the proof for the general σ-finite case.]

Exercise 1.1.9. Define a premeasure on finite unions of boxes $\prod_{i=1}^{n} U_i$ with $U_i \in \mathcal{X}_i$ by $\mu(\prod_{i=1}^{n} U_i) := \prod_{i=1}^{n} \mu_i(U_i)$ and extending to unions by decomposing them into disjoint boxes. We may then apply the Carathéodory extension theorem. (See 245A Proposition 1.7.11.)

Exercise 1.1.10. [I'm skipping this exercise.]

Exercise 1.1.11. The sequence of functions

$$|f| 1_{\{x \in E : |f(x)| > 1\}} \geq |f| 1_{\{x \in E : |f(x)| > 2\}} \geq \cdots$$

converges pointwise a.e. to the zero function, since f is absolutely integrable. Thus we have

$$\lim_{n\to\infty} \int_X |f| 1_{\{x \in E : |f(x)| > n\}} \, d\mu = \int_X \lim_{n\to\infty} |f| 1_{\{x \in E : |f(x)| > n\}} \, d\mu = 0$$

by dominated convergence, and so we may choose large λ for which $\int_{x\in E:|f(x)|>\lambda} |f|\,d\mu \leq \epsilon/2$. It follows that

$$\begin{aligned}\int_E |f|\,d\mu &= \int_{x\in E:|f(x)|\leq\lambda} |f|\,d\mu + \int_{x\in E:|f(x)|>\lambda} |f|\,d\mu \\ &\leq \lambda\mu(E) + \epsilon/2 \\ &\leq \epsilon\end{aligned}$$

whenever $\mu(E) \leq \epsilon/2\lambda$.

Make use of time, let not advantage slip;
Beauty within itself should not be wasted:
Fair flowers that are not gather'd in their prime
Rot and consume themselves in little time.

— William Shakespeare, *Venus and Adonis* (1593)

1.2. Signed measures and the Radon–Nikodym–Lebesgue theorem

Observe due measure, for right timing is in all things the most important factor.

— Hesiod, *Works and Days* (c. 700 B.C.)

Exercise 1.2.1. We first prove that m_f is an unsigned measure. We have

$$m_f(\varnothing) = \int_X 1_\varnothing f\, dm = \int_X 0\, dm = 0.$$

Given disjoint $E_1, E_2, \dots$, we have

$$\begin{aligned} m_f\Big(\bigcup_{n=1}^\infty E_n\Big) &= \int_X 1_{\bigcup_{n=1}^\infty E_n} f\, dm \\ &= \int_X \sum_{n=1}^\infty 1_{E_n} f\, dm \\ &= \sum_{n=1}^\infty \int_X 1_{E_n} f\, dm \\ &= \sum_{n=1}^\infty m_f(E_n), \end{aligned}$$

where we used monotone convergence for series (Theorem 1.1.21) to swap the sum and integral.

Suppose $g \colon X \to [0, +\infty]$ is a simple unsigned function taking values in $\{a_1, \dots, a_n\}$. We write $g = \sum_{i=1}^n a_i 1_{g^{-1}(\{a_i\})}$, and compute

$$\begin{aligned} \int_X g\, dm_f &= \sum_{i=1}^n a_i m_f(g^{-1}(\{a_i\})) \\ &= \sum_{i=1}^n a_i \int_X 1_{g^{-1}(\{a_i\})} f\, dm \\ &= \int_X \sum_{i=1}^n a_i 1_{g^{-1}(\{a_i\})} f\, dm \\ &= \int_X gf\, dm. \end{aligned}$$

Since every unsigned measurable function is the pointwise limit of an increasing sequence of unsigned simple functions[1], the result for general g follows from monotone convergence (Theorem 1.1.21).

[1] This result is occasionally called the *Sombrero lemma* due to the construction of the sequence of functions involved. See René L. Schilling, *Measures, Integrals and Martingales* 2e., Theorem 8.8.

Exercise 1.2.2. If $f = g$ m-a.e., then $m_f(E) = \int_E f = \int_E g = m_g(E)$, as the Lebesgue integrals are equal for a.e. equal functions.

Conversely, we prove that $m_f = m_g$ implies that $f = g$ m-a.e.. We first consider the case where $m(X) < \infty$. Suppose contrapositively that $f \neq g$ m-a.e.. Then, without loss of generality, there exists a set E of positive finite measure such that $f > g$ on E. We consider two cases.

Case 1: $\int_E f, \int_E g < \infty$. In this case, f and g must be finite m-a.e., and thus we may safely consider the function $f - g$ on E, which is unsigned measurable as $f > g$ on E by hypothesis. Therefore, we have $\int_E f - g \leq \int_E f < \infty$ and

$$\int_E f = \int_E f - g + \int_E g.$$

Since $f - g > 0$ on E, we have $\int_E f - g > 0$, and so $m_f(E) > m_g(E)$ as desired.

Case 2: $\int_E f = \infty$. If $\int_E g < \infty$, there is nothing to prove, so assume that $\int_E f = \infty = \int_E g$. Since $f > g$, we see that g must be finite everywhere. Apply monotone convergence to the sequence $g1_{\{x\in E:g(x)\leq 1\}} \leq g1_{\{x\in E:g(x)\leq 2\}} \leq \dots$ to obtain the identity

$$\int_E g = \lim_{N\to\infty} \int_{x\in E:g(x)\leq N} g.$$

It follows that there exists N such that $m(\{x \in E : g(x) \leq N\}) > 0$. Let $E' := \{x \in E : g(x) \leq N\}$. Since $\int_{E'} g \leq Nm(E') < \infty$, we are left to consider $\int_{E'} f$. If $\int_{E'} f = \infty$, we are done. Otherwise, we have $\int_{E'} f < \infty$, and we are left with case 1.

This concludes the proof for the finite measure case.

Now suppose that m is σ-finite. Write $X = \bigcup_{n=1}^{\infty} X_n$, with X_n disjoint and $m(X_n) < \infty$. Then, once again, if $f > g$ on E with $m(E) > 0$ (possibly infinite this time), then we may consider the finite measure sets $E \cap X_n$. At least one of these sets $E \cap X_n$ is non-empty, with $m(E \cap X_n) < \infty$. Thus we may apply the finite measure argument above to obtain a set $E' \subset E \cap X_n$ on which $\int_{E'} f > \int_{E'} g$ as needed.

Finally, we give a counterexample when μ fails to be σ-finite. Consider the measurable space $(\mathbf{N}, 2^{\mathbf{N}})$ equipped with the measure $\mu(E) = +\infty \cdot [E \text{ is non-empty}]$. That is, μ gives all non-empty sets infinite measure. Then, setting $f = 1_{\mathbf{N}}$ and $g = 2 \cdot 1_{\mathbf{N}}$, we see that $\int_X 1_\varnothing f \, d\mu = 0 = \int_X 1_\varnothing g \, d\mu$, and that

$$\int_X 1_E f \, d\mu = +\infty = \int_X 1_E g \, d\mu$$

for all non-empty $E \in 2^{\mathbf{N}}$ (this idea works with a singleton set, but I found $\mathbf{N}$ more comforting).

Exercise 1.2.3. To say that μ has a continuous Radon–Nikodym derivative $d\mu/dm$ is to say that there exists a continuous function $f = d\mu/dm$ such that $\mu = m_f$. We thus compute

$$\mu([0, x]) = m_f([0, x]) = \int_{[0,x]} f \, dm.$$

By the fundamental theorem of calculus, we conclude that

$$\frac{d}{dx}\mu([0, x]) = f(x) = \frac{d\mu}{dm}(x)$$

for all $x \in [0, +\infty)$.

Exercise 1.2.4. Let $\mu\colon X \to [0, +\infty]$ be a measure on X. We would like to write $\mu = \#_f$ for some function $f\colon X \to [0, +\infty]$. Expanding the definitions, we are looking for some f such that

$$\mu(E) = \#_f(E) = \int_E f \, d\# = \sum_{x \in E} f(x).$$

Thus we conclude that the function f, defined by $f(x) := \mu(\{x\})$, is indeed the Radon–Nikodym derivative $d\mu/d\#$ of μ with respect to #.

Remark. If a measure μ on X is differentiable with respect to the Dirac measure δ_x with Radon–Nikodym derivative $d\mu/d\delta_x = f$, then we must have $\mu(E) = (\delta_x)_f(E) = \int_E f \, d\delta_x = f(x)\delta_x(E)$. Since the Radon–Nikodym derivative is defined up to δ_x-a.e. equivalence (which means that $f = g$ δ_x-a.e. iff $f(x) = g(x)$), we see that the only measures differentiable with respect to δ_x are its scalar multiples.

Exercise 1.2.5. Let $\mu = \mu \lfloor_{X_+} - \mu \lfloor_{X_-} = \mu_+ - \mu_-$ be as obtained from the Hahn decomposition theorem, and suppose $\mu = \nu_+ - \nu_-$ is another decomposition such that ν_+ and ν_- are mutually singular unsigned measures. Since ν_+ and ν_- are mutually singular, we may write X as a disjoint union $X = Y_+ \cup Y_-$ such that ν_+ is supported on Y_+ and ν_- is supported on Y_-. Then we may write X as the disjoint union of four sets, namely

$$X = (X_+ \cap Y_+) \cup (X_+ \cap Y_-) \cup (X_- \cap Y_+) \cup (X_- \cap Y_-).$$

On $X_+ \cap Y_+$, we have $\mu_- = \nu_- = 0$, and so

$$\mu_+ = \mu_+ - \mu_- = \nu_+ - \nu_- = \nu_+;$$

consequently $\mu_- = \nu_-$. On $X_+ \cap Y_-$, we have $\mu_- = \nu_+ = 0$, and so

$$\mu_+ = \nu_+ - \nu_- + \mu_+ = -\nu_-.$$

Since μ_+ and ν_- are unsigned, it follows that $\mu_+ = \nu_- = 0$, and so $\mu_+ = \nu_+$ as needed. The remaining cases are handled similarly.

Exercise 1.2.6. We first verify that $|\mu|$ is an unsigned measure. Since μ_+ and μ_- are unsigned, we see that $|\mu| = \mu_+ + \mu_-$ is unsigned as well. Given disjoint $E_1, E_2, \ldots \subset X$, we may compute

$$\begin{aligned} |\mu|\Big(\bigcup_{n=1}^\infty E_n\Big) &= \mu_+\Big(\bigcup_{n=1}^\infty E_n\Big) + \mu_-\Big(\bigcup_{n=1}^\infty E_n\Big) \\ &= \sum_{n=1}^\infty \mu_+(E_n) + \sum_{n=1}^\infty \mu_-(E_n) \\ &= \sum_{n=1}^\infty |\mu|(E_n), \end{aligned}$$

where the last equality is justified by the absolute convergence of both series.

Let ν be an unsigned measure such that $-\nu \leq \mu \leq \nu$, or

$$-\nu_+ + \nu_- \leq \mu_+ - \mu_- \leq \nu_+ - \nu_-.$$

Our goal is to prove that $|\mu| \leq \nu$, or $\mu_+ + \mu_- \leq \nu_+ - \nu_-$. Applying Hahn decomposition to μ, we get $X = X_+ \cup X_-$. Similarly, applying

Hahn decomposition to ν gives $X = Y_+ \cup Y_-$. We may thus write X as a disjoint union

$$X = (X_+ \cap Y_+) \cup (X_+ \cap Y_-) \cup (X_- \cap Y_+) \cup (X_- \cap Y_-).$$

On $X_+ \cap Y_+$, we have $\mu_- = \nu_- = 0$. Thus

$$\mu_+ + \mu_- = \mu_+ - \mu_- \leq \nu_+ - \nu_-$$

as needed. The remaining three cases are handled similarly.

Now we prove that

$$|\mu|(E) = \sup \sum_{n=1}^{\infty} |\mu(E_n)|,$$

where the supremum is taken over all partitions $(E_n)_{n=1}^{\infty}$ of E. Since

$$-\sup \sum_{n=1}^{\infty} |\mu(E_n)| \leq \mu(E) \leq \sup \sum_{n=1}^{\infty} |\mu(E_n)|,$$

earlier arguments imply that $|\mu|(E) \leq \sup \sum_{n=1}^{\infty} |\mu(E_n)|$. Conversely, since $-|\mu| \leq \mu \leq |\mu|$ means that $|\mu(E)| \leq |\mu|(E)$, we have

$$\sum_{n=1}^{\infty} |\mu(E_n)| \leq \sum_{n=1}^{\infty} |\mu|(E_n) = |\mu|\Big(\bigcup_{n=1}^{\infty} E_n\Big) = |\mu|(E)$$

for any partition $(E_n)_{n=1}^{\infty}$ of E, and so we conclude that

$$|\mu|(E) = \sup \sum_{n=1}^{\infty} |\mu(E_n)|$$

as needed.

Exercise 1.2.7. We prove that the following are equivalent:

(i) $\mu(E)$ is finite for every $E \subset X$.

(ii) $|\mu|$ is a finite unsigned measure.

(iii) μ_+ and μ_- are finite unsigned measures.

Claim (i) implies (ii). Indeed, if $\mu(E)$ is finite, then $\mu_+(E) - \mu_-(E)$ is finite. Since the quantities cannot both be infinite, they must both be finite, and so $|\mu|(E) = \mu_+(E) + \mu_-(E)$ is finite as well.

Claim (ii) implies (iii), since

$$\mu_+(E) \leq |\mu|(E) < \infty;$$

similarly for μ_-.

Finally, (iii) implies (i) as

$$\mu(E) = \mu_+(E) - \mu_-(E) < \infty.$$

Remark (Proof of Theorem 1.2.4). [Folland 2e, Lemma 3.7] In the last paragraph of the proof of Theorem 1.2.4, it is shown that $\mu_s \perp m$. Here are some details:

We prove that either $\mu_s \perp m$, or there exist $\epsilon > 0$ and $E \in \mathcal{X}$ such that $m(E) > 0$ and $\mu_s \geq \epsilon m$ on E (that is, E is a totally positive set for $\mu_s - \epsilon m$).

Indeed, let $X = X^n_+ \cup X^n_-$ be a Hahn decomposition for $\mu_s - n^{-1}m$, and let $X_+ := \bigcup_{n=1}^\infty X^n_+$ and $X_- := \bigcap_{n=1}^\infty X^n_- = X \setminus X^+$. Then X_- is a totally negative set for $\mu_s - n^{-1}m$ for all n; i.e., $0 \leq \mu_s(X_-) \leq n^{-1}m(X_-)$ for all n, so $\mu_s(X_-) = 0$. If $m(X_+) = 0$, then $\mu_s \perp m$. Otherwise, if $m(X_+) > 0$, then $m(X^n_+) > 0$ for some n, and so $X^n_+ =: E$ is a totally positive set for $\mu_s - n^{-1}m$.

Exercise 1.2.8. [Folland 2e, Theorem 3.8; Math.SE answer 3713882] (I'm still a bit sketchy on this solution.) Suppose μ, m are σ-finite. Then we may write $X = \bigcup_{n=1}^\infty X_n$ with $\mu(X_n) < \infty$, $m(X_n) < \infty$, and X_n disjoint. Defining $\mu_n(E) := \mu(E \cap X_n)$ and $m_n(E) := m(E \cap X_n)$, we may apply the result for the finite measure case to obtain decompositions

$$\mu_n = (m_n)_{f_n} + (\mu_n)_s$$

with $(\mu_n)_s \perp m_n$. Let $f := \sum_n f_n$ and $\mu_s := \sum_n (\mu_n)_s$. We may assume that $f_n = 0$ on $X \setminus X_n$, so that

$$\begin{aligned}
\sum_n (m_n)_{f_n}(E) &= \sum_n \int_E f_n \, dm_n \\
&= \sum_n \int_E f_n \, dm \\
&= \int_E \sum_n f_n \, dm \\
&= \int_E f \, dm \\
&= m_f(E).
\end{aligned}$$

Thus, we have

$$\mu = \sum_n \mu_n = \sum_n (m_n)_{f_n} + \sum_n (\mu_n)_s = m_f + \mu_s.$$

Finally, we prove that $\mu_s \perp m$. Since $(\mu_n)_s \perp m_n$, we may write $X = A_n \cup B_n$ with A_n, B_n disjoint such that $(\mu_n)_s$ is null outside A_n and m_n is null outside B_n. Then, setting $\tilde{A}_n := A_n \cap X_n$ and $\tilde{B}_n := B_n \cap X_n$, we may define $A := \bigcup_n \tilde{A}_n$ and $B := \bigcup_n \tilde{B}_n$, so that $X = A \cup B$ with A, B disjoint. Since $\mu_s = \sum_n (\mu_n)_s$ is null outside A and $m = \sum_n m_n$ is null outside B, we conclude that $\mu_s \perp m$ as needed.

Exercise 3.9 from Folland 2e. *Suppose (ν_n) is a sequence of unsigned measures. If $\nu_n \perp \mu$ for all n, then $\sum_n \nu_n \perp \mu$; and if $\nu_n \ll \mu$ for all n, then $\sum_n \nu_n \ll \mu$.*

Say ν_n is supported on X_n, so that μ is supported on $X \setminus X_n$. Then $\sum_n \nu_n$ is supported on $\bigcup_n X_n$, and μ is supported on $\bigcap_n (X \setminus X_n) = X \setminus \bigcup_n X_n$, so that $\sum_n \nu_n \perp \mu$. Suppose $\nu_n(E) = 0$ whenever $\mu(E) = 0$. Then $\sum_n \nu_n(E) = 0$ whenever $\mu(E) = 0$, and so $\sum_n \nu_n \ll \mu$.

Exercise 1.2.9. Let m be an unsigned σ-finite measure. As before, by Hahn decomposition, we may assume that μ is an *unsigned* σ-finite measure. Suppose every point is measurable, and that $m(\{x\}) = 0$ for all $x \in X$. (We say that m is *continuous*.) By the Lebesgue decomposition theorem, we may write $\mu = \mu_{\text{ac}} + \mu_{\text{s}}$ uniquely, with $\mu_{\text{ac}} \ll m$ and $\mu_{\text{s}} \perp m$. We will further decompose

$$\mu_{\text{s}} = \mu_{\text{sc}} + \mu_{\text{pp}},$$

where μ_{pp} is supported on an at most countable set, and where μ_{sc} is continuous with $\mu_{\text{sc}} \perp m$. The natural idea is to define the set

$$E := \{x \in X : \mu_{\text{s}}(\{x\}) > 0\}.$$

Let $\mu_{\text{sc}} := \mu_{\text{s}}\lfloor_{X \setminus E}$ and $\mu_{\text{pp}} := \mu_{\text{s}}\lfloor_{E}$. Then we must show:

(1) E is at most countable.

(2) $\mu_{\text{sc}}(\{x\}) = 0$ for all $x \in X$.

(3) $\mu_{\text{sc}} \perp m$.

We first prove (1). Suppose for contradiction that E is uncountable. Since $\mu_{\text{s}} \leq \mu$ and μ is σ-finite, it follows that μ_{s} is σ-finite as well. Thus we may write $X = \bigcup_{n=1}^{\infty} X_n$ such that $\mu_{\text{s}}(X_n) < \infty$, with X_n disjoint. Then, $E = \bigcup_{n=1}^{\infty}(E \cap X_n)$, with $\mu_{\text{s}}(E \cap X_n) < \infty$ and $E \cap X_n$ disjoint. Since the countable union of countable sets is countable, there exists n such that $E \cap X_n$ is uncountable. Define

$$E_{n,k} := \left\{x \in E \cap X_n : \frac{1}{k} \leq \mu_s(\{x\}) < \frac{1}{k-1}\right\}$$

for $k \geq 2$, with $E_{n,1} := \{x \in E \cap X_n : \mu_s(\{x\}) \geq 1\}$. Then $E \cap X_n = \bigcup_{k=1}^{\infty} E_{n,k}$ with $E_{n,k}$ disjoint, and so $E_{n,k}$ is uncountable for some k. Taking a countable subset $S \subset E_{n,k}$, we see that

$$\mu(E \cap X_n) \geq \mu(E_{n,k}) \geq \sum_{j=1}^{\infty} \frac{1}{k} = +\infty,$$

contradicting the finiteness of $\mu(E \cap X_n)$.

Claim (2) holds as μ_{sc} is supported on $X \setminus E$, and all positive measure singletons are in E by definition.

Claim (3) follows from the fact that $\mu_{\text{s}} \perp m$. Indeed, $\mu_{\text{sc}} \leq \mu_{\text{s}}$, which implies that the support of μ_{sc} is a subset of the support of μ_{s}. This completes the proof.

Remark (Absolute continuity). [C. Heil, *Introduction to Real Analysis*, Problem 6.1.9] Using the definition in the text, we can prove that a function $f \colon I \to \mathbf{R}$ is absolutely continuous if and only if, for every $\epsilon > 0$, there exists $\delta > 0$ such that $\sum_{i=1}^{\infty} |f(y_i) - f(x_i)| \leq \epsilon$ whenever $[x_1, y_1], \ldots$ is a family of *countably many* disjoint intervals in I of total length at most δ.

Indeed, given $\epsilon > 0$ and a countably infinite family $[x_1, y_1], \ldots,$ we choose $\delta > 0$ as in the finite case. If $\sum_{i=1}^{\infty}(y_i - x_i) < \delta$, then $\sum_{i=1}^{n}(y_i - x_i) < \delta$ for all n, and therefore $\sum_{i=1}^{n} |f(y_i) - f(x_i)| \leq \epsilon$

for all n. Thus we conclude that $\sum_{i=1}^{\infty} |f(y_i) - f(x_i)| \le \epsilon$ as needed. The converse follows from setting all sufficiently large intervals to be empty.

Exercise 1.2.10. (i) Suppose μ is continuous. We first prove that $x \mapsto \mu([0,x])$ is right-continuous. This does not require the continuity hypothesis for μ (it is a general property of *cumulative distribution functions*). Indeed,

$$\lim_{h\downarrow 0} \mu([0,x+h]) - \mu([0,x]) = \lim_{h\downarrow 0} \mu((x,x+h]) \le \mu((x,x+1/n))$$

for all $n \ge 1$, and so

$$\begin{aligned}\lim_{h\downarrow 0} \mu((x,x+h]) &\le \lim_{n\to\infty} \mu((x,x+1/n)) \\ &= \mu\Big(\bigcap_{n=1}^{\infty} (x,x+1/n)\Big) \\ &= \mu(\varnothing) \\ &= 0\end{aligned}$$

by dominated convergence for sets.

Now we prove left-continuity. We must prove that

$$\lim_{h\downarrow 0} \mu([0,x-h]) - \mu([0,x]) = 0,$$

or equivalently, that

$$\lim_{h\downarrow 0} \mu((x-h,x]) = 0.$$

Arguing as before, we see that

$$\begin{aligned}\lim_{h\downarrow 0} \mu((x-h,x]) &\le \lim_{n\to\infty} \mu((x-1/n,x]) \\ &= \mu\Big(\bigcap_{n=1}^{\infty} (x-1/n,x]\Big) \\ &= \mu(\{x\}) \\ &= 0\end{aligned}$$

as needed.

Conversely, suppose $x \mapsto \mu([0,x])$ is continuous. Fix $x \in [0,+\infty]$. We prove that $\mu(\{x\}) \le \epsilon$ for all $\epsilon > 0$. By continuity,

$$\lim_{h\downarrow 0} \mu([0,x-h]) = \mu([0,x]),$$

so that

$$\lim_{h\downarrow 0} \mu((x-h,x]) = 0.$$

Thus, for small h we have

$$\mu(\{x\}) \le \mu((x-h,x]) \le \epsilon$$

as needed.

(ii) Let $\epsilon > 0$. If $\mu \ll m$, then the Radon–Nikodym theorem tells us that there exists $\delta > 0$ such that $|\mu(E)| < \epsilon$ whenever $m(E) \le \delta$. Suppose $[x_1, y_1], \dots, [x_n, y_n]$ are disjoint intervals in $[0, +\infty]$ of total length at most δ, so that $m(\bigcup_{i=1}^n (x_i, y_i]) \le \delta$. Then

$$\sum_{i=1}^n \big|\mu([0, y_i]) - \mu([0, x_i])\big| = \sum_{i=1}^n \mu((x_i, y_i]) = \mu\Big(\bigcup_{i=1}^n (x_i, y_i]\Big) < \epsilon,$$

proving that $x \mapsto \mu([0, x])$ is absolutely continuous. Conversely, suppose $x \mapsto \mu([0, x])$ is absolutely continuous. By the remarks above, there exists $\delta > 0$ such that $\mu(\bigcup_{i=1}^\infty (x_i, y_i]) < \epsilon$ whenever $[x_1, y_1], \dots$ is a countable family of disjoint intervals in $[0, +\infty]$ of total length at most δ. Suppose $m(E) = 0$. Then outer regularity of Lebesgue measure implies that we may cover E with an open set U of m-measure at most δ. Open sets in $\mathbf{R}$ can be written as countable unions of open intervals; thus we may write $U = \bigcup_{i=1}^\infty (a_i, b_i)$ so as to conclude that $\mu(E) \le \mu(U) < \epsilon$. Since ϵ is arbitrary, we conclude that $\mu(E) = 0$ as needed.

Exercise 1.2.11. [I'm skipping this exercise.]

I saw myself sitting in the crotch of this fig-tree, starving to death,
just because I couldn't make up my mind which of the figs I would choose.
I wanted each and every one of them,
but choosing one meant losing all the rest, and,
as I sat there, unable to decide, the figs began to wrinkle and go black, and,
one by one, they plopped to the ground at my feet.

— Sylvia Plath, *The Bell Jar* (1963)

1.3. L^p spaces

... if one were to refuse to have direct, geometric, intuitive insights,
if one were reduced to pure logic, which does not permit a choice among
every thing that is exact, one would hardly think of many questions,
and certain notions ... would escape us completely.

— HENRI LEBESGUE, *Sur le développement de la notion d'intégrale* (1926)

Exercise 1.3.1. We are given the space $L^p(X, \mathcal{X}, \mu)$ together with its completion $L^p(X, \overline{\mathcal{X}}, \overline{\mu})$. Every function $\overline{f}\colon X \to \mathbf{C}$ that is measurable with respect to $(X, \overline{\mathcal{X}}, \overline{\mu})$ can be associated with a function $f\colon X \to \mathbf{C}$ that is measurable with respect to $(X, \mathcal{X}, \mu)$, such that

$$\overline{\mu}(\{x \in X : f(x) \neq \overline{f}(x)\}) = 0.$$

It suffices to prove this for simple functions, as a measurable function is the supremum of a sequence of simple functions. Suppose $\overline{E_i}$ is $\overline{\mathcal{X}}$-measurable. Then, by definition of the completion, $\overline{E_i}$ must differ from an $\mathcal{X}$-measurable set E_i by a sub-null set, so that $\overline{\mu}(\overline{E_i}) = \mu(E_i)$. Thus

$$\int_X \sum_{i=1}^n c_i 1_{\overline{E_i}}\, d\overline{\mu} = \sum_{i=1}^n c_i \overline{\mu}(\overline{E_i}) = \sum_{i=1}^n c_i \mu(E_i) = \int_X \sum_{i=1}^n c_i 1_{E_i}\, d\mu.$$

Exercise 1.3.2. (i) We would like to argue that

$$\begin{aligned} \|f+g\|_{L^p}^p &= \int_X |f(x)+g(x)|^p \, d\mu \\ &\overset{?}{\leq} \int_X |f(x)|^p + |g(x)|^p \, d\mu \\ &= \|f\|_{L^p}^p + \|g\|_{L^p}^p. \end{aligned}$$

Thus, given $x \in X$, it suffices to prove that

$$|f(x)+g(x)|^p \leq |f(x)|^p + |g(x)|^p, \qquad 0 < p < 1 \qquad (*)$$

whenever $f(x)$ and $g(x)$ are both non-zero. This in turn follows from the real inequality

$$(1+t)^p \leq 1 + t^p, \qquad t \geq 0,\ \ 0 < p < 1,$$

as then, for $\alpha \in \mathbf{C}$, the complex triangle inequality implies that

$$|1+\alpha|^p \leq (1+|\alpha|)^p \leq 1 + |\alpha|^p;$$

the inequality $(*)$ then follows by setting $\alpha = f(x)/g(x)$. Since the function $h(t) := 1 + t^p - (1+t)^p$ for $t \geq 0$ is such that $h(0) = 0$ and $h'(t) = pt^{p-1} - p(1+t)^{p-1} = p(1/t^{1-p} - 1/(1+t)^{1-p}) \geq 0$, it must be a non-decreasing function, and thus the result follows.

(ii) We emulate the proof of Lemma 1.3.3(iii), except this time the function $x \mapsto |x|^p$ for $x > 0$ is *concave* as we have $0 < p < 1$. As before, by non-degeneracy we may take both $\|f\|_{L^p}$ and $\|g\|_{L^p}$ to be non-zero. By homogeneity we normalize $\|f\|_{L^p} + \|g\|_{L^p} = 1$, and by homogeneity again we write $f = (1-\theta)F$ and $g = \theta G$ for some $0 < \theta < 1$ and $F, G \in L^p$ with $\|F\|_{L^p} = \|G\|_{L^p} = 1$. Our task is then to show that

$$\int_X \big((1-\theta)F(x) + \theta G(x)\big)^p \, d\mu \geq 1. \qquad (*)$$

Since the function $x \mapsto x^p$ is concave for $x > 0$ and $0 < p < 1$, we have

$$\big((1-\theta)F(x) + \theta G(x)\big)^p \geq (1-\theta)F(x)^p + \theta G(x)^p.$$

Together with the normalizations of $\|F\|_{L^p}$ and $\|G\|_{L^p}$, this implies $(*)$ as desired.

(iii) By (i), we have $\|f+g\|_{L^p} \leq (\|f\|_{L^p}^p + \|g\|_{L^p}^p)^{1/p}$. Since $x \mapsto x^{1/p}$ is convex for $x > 0$ and $0 < p < 1$, we have

$$\Big(\frac{1}{2}\|f\|_{L^p}^p + \frac{1}{2}\|g\|_{L^p}^p\Big)^{1/p} \leq \frac{1}{2}\|f\|_{L^p} + \frac{1}{2}\|g\|_{L^p}.$$

It follows that

$$\|f+g\|_{L^p} \leq 2^{1/p-1}(\|f\|_{L^p} + \|g\|_{L^p}).$$

This constant is in fact best possible, since we may take, say, $f = 1_{[0,1]}$ and $g = 1_{[1,2]}$ to get

$$\|f+g\|_{L^p} = 2^{1/p} = 2^{1/p-1}(1+1) = 2^{1/p-1}(\|f\|_{L^p} + \|g\|_{L^p}).$$

(iv) First suppose $0 < p < 1$. Since $x \mapsto x^p$ is non-linear, the only way equality can occur in Jensen's inequality

$$\big((1-\theta)F(x) + \theta G(x)\big)^p \geq (1-\theta)F(x)^p + \theta G(x)^p$$

is when $F(x) = G(x)$. This implies that $f = cg$ for some $c > 0$. The case for $p > 1$ is analogous.

When $p = 1$, the identity becomes

$$\int_X |f(x) + g(x)| \, d\mu = \int_X |f(x)| \, d\mu + \int_X |g(x)| \, d\mu,$$

which holds for all non-negative measurable functions f and g by linearity of the integral.

Exercise 1.3.3. Let $\|\cdot\|$ be a norm, and let $v, w \in \{x \in V : \|x\| \leq 1\}$. Then, given $0 < t < 1$, homogeneity and the triangle equality imply that

$$\|tv + (1-t)w\| \leq |t|\|v\| + |1-t|\|w\| \leq t + (1-t) = 1,$$

so that the line joining v and w is contained in the closed unit ball. Conversely, suppose that the closed unit ball is convex. Then, given $v, w \in V$, we must prove the triangle inequality. By non-degeneracy, we may assume both vectors are non-zero. By homogeneity, we may assume that $\|v\| + \|w\| = 1/2$. By homogeneity again, we can write $v = (1-\theta)v'$ and $w = \theta w'$ for some $0 < \theta < 1$ and $v', w' \in V$ with $\|v'\| = \|w'\| = 1/2$. Convexity then implies that

$$\|v + w\| = \|(1-\theta)v' + \theta w'\| \leq (1-\theta)\|v'\| + \theta\|w'\| = \|v\| + \|w\|,$$

as desired. The proofs for the open unit ball are analogous.

Exercise 1.3.4. Note that $\operatorname{supp} f = \operatorname{supp} |f|^p$. Markov's inequality implies that

$$\mu\Big(\Big\{x \in X : |f(x)|^p \geq \frac{1}{n}\Big\}\Big) \leq n \int_X |f(x)|^p \, d\mu < \infty.$$

Thus

$$\operatorname{supp} f = \bigcup_{n=1}^{\infty} \Big\{x \in X : |f(x)|^p \geq \frac{1}{n}\Big\}$$

is σ-finite.

Exercise 1.3.5.

(i) [I could not solve this. This solution is an elaboration of `https://math.stackexchange.com/a/242792/` for my own understanding.]

Let $0 < \delta < \|f\|_{L^\infty}$, and let $S_\delta := \{x \in X : |f(x)| \geq \|f\|_{L^\infty} - \delta\}$. By definition of $\|\cdot\|_{L^\infty}$, we have $\mu(S_\delta) > 0$. We compute

$$\|f\|_{L^p} \geq \Big(\int_{S_\delta} (\|f\|_{L^\infty} - \delta)^p \, d\mu\Big)^{1/p} = (\|f\|_{L^\infty} - \delta)\mu(S_\delta)^{1/p} \qquad (*)$$

for $0 < p < \infty$. Setting $p = p_0$, we see that

$$(\|f\|_{L^\infty} - \delta)\mu(S_\delta)^{1/p_0} \leq \|f\|_{L^{p_0}} < \infty,$$

so that $\mu(S_\delta) < \infty$. Taking the limit inferior as $p \to \infty$ of $(*)$, we thus have

$$\liminf_{p\to\infty} \|f\|_{L^p} \geq \|f\|_{L^\infty}.$$

Conversely, since $|f(x)| \leq \|f\|_{L^\infty}$ for almost every x, we have

$$\begin{aligned}
\|f\|_{L^p} &= \Big(\int_X |f(x)|^{p-p_0} |f(x)|^{p_0} \, d\mu\Big)^{1/p} \\
&\leq \Big(\int_X \|f\|_{L^\infty}^{p-p_0} |f(x)|^{p_0} \, d\mu\Big)^{1/p} \\
&= \|f\|_{L^\infty}^{(p-p_0)/p} \Big(\int_X |f(x)|^{p_0} \, d\mu\Big)^{1/p} \\
&= \|f\|_{L^\infty}^{(p-p_0)/p} \|f\|_{L^{p_0}}^{p_0/p}
\end{aligned}$$

whenever $p > p_0$. Taking the limit superior as $p \to \infty$, we conclude that

$$\limsup_{p\to\infty} \|f\|_{L^p} \leq \|f\|_{L^\infty}.$$

Therefore, the limit $\lim_{p\to\infty} \|f\|_{L^p}$ exists and is equal to $\|f\|_{L^\infty}$.

(ii) The argument is a modification of (i), except this time we use sets of the form $S_N := \{x \in X : |f(x)| \geq N\}$. We also handle the case $\mu(S_N) = +\infty$ directly, and we do not need the limit superior case.

Exercise 1.3.6. These are routine verifications. We first verify that the function d is a metric:

- (Non-degeneracy) By non-degeneracy of $\|\cdot\|$, we have $d(f, g) = 0$ iff $\|f - g\| = 0$ iff $f - g = 0$ iff $f = g$.

- (Symmetry) By homogeneity of $\|\cdot\|$, we have
$$d(f,g) = \|f-g\| = |-1|\|g-f\| = d(g,f).$$
- (Triangle inequality) By the triangle inequality for $\|\cdot\|$, we have
$$d(f,h) = \|f-h\| \le \|f-g\| + \|g-h\| = d(f,g) + d(g,h).$$

This metric d satisfies:

- (Translation-invariance) We have
$$d(f+h, g+h) = \|(f+h)-(g+h)\| = \|f-g\| = d(f,g).$$
- (Homogeneity) By homogeneity of $\|\cdot\|$, we have
$$d(cf,cg) = \|cf-cg\| = \|c(f-g)\| = |c|\|f-g\| = |c|d(f,g).$$

Conversely, given a translation-invariant homogeneous metric d, we may define a function $\|\cdot\| \colon V \to [0,+\infty)$ by $\|f\| := d(0,f)$. We verify that this function $\|\cdot\|$ is a norm:

- (Non-degeneracy) By the non-degeneracy of d, we have $\|f\| = d(0,f) = 0$ iff $f = 0$.
- (Homogeneity) By homogeneity of d, we have
$$\|cf\| = d(0,cf) = |c|d(0,f) = |c|\|f\|.$$
- (Triangle inequality) By the triangle inequality for d, and by the translation-invariance of d, we have
$$\begin{aligned}\|f+g\| &= d(0,f+g)\\ &\le d(0,f) + d(f,f+g)\\ &= d(0,f) + d(0,g)\\ &= \|f\| + \|g\|.\end{aligned}$$

We may establish analogous claims relating quasi-norms and quasi-metrics, as well as seminorms and semimetrics.

Exercise 1.3.7. Suppose the series $\sum_{j=1}^{\infty} f_j$ converges absolutely, so that $\sum_{j=1}^{\infty} \|f_j\| < \infty$. We claim that $(\sum_{j=1}^{n} f_j)_{n=1}^{\infty}$ is a Cauchy sequence. Let $\epsilon > 0$. Then there exists N such that $\sum_{j=N}^{\infty} \|f_j\| < \epsilon$. Thus, given $m, n \ge N$, we have by repeated applications of the triangle inequality
$$\begin{aligned}\Big\|\sum_{j=m}^{n} f_j\Big\| &\le \sum_{j=m}^{n} \|f_j\|\\ &\le \sum_{j=N}^{\infty} \|f_j\|\\ &< \epsilon.\end{aligned}$$

Therefore $(\sum_{j=1}^n f_j)_{n=1}^\infty$ is Cauchy and converges to a limit f, which must be equal to $\sum_{j=1}^\infty f_j$ by definition of summation. Thus $\sum_{j=1}^\infty f_j$ is conditionally convergent as needed.

Conversely, suppose that absolute convergence implies conditional convergence, and let $(f_j)_{j=1}^\infty$ be a Cauchy sequence. Choose a sequence of integers $N_1 < N_2 < N_3 < \dots$ such that $\|f_m - f_n\| \le \epsilon/2^k$ whenever $m, n \ge N_k$. Then $\sum_{j=k}^\infty \|f_{N_j} - f_{N_{j-1}}\| \le \epsilon/2^{k-1}$, where $k \ge 2$. Therefore the series $\sum_{j=2}^\infty \|f_{N_j} - f_{N_{j-1}}\|$ converges, and by hypothesis the series $\sum_{j=2}^\infty (f_{N_j} - f_{N_{j-1}})$ converges as well. Since

$$\lim_{k\to\infty} \sum_{j=2}^k (f_{N_j} - f_{N_{j-1}}) = \lim_{k\to\infty} (f_{N_k} - f_{N_1}) = \lim_{k\to\infty} f_{N_k} - f_{N_1},$$

we see that the limit $\lim_{k\to\infty} f_{N_k}$ exists. Thus we have a convergent subsequence of a Cauchy sequence, which implies that the original sequence converges as desired.

Remark. Here is an equivalent formulation of convergence in L^p norm that is useful for understanding the last sentence of the proof of Proposition 1.3.7. Let $1 \le p < \infty$. Given a sequence $(f_n)_{n=1}^\infty$ of L^p functions together with an L^p function f, we have

$$\lim_{n\to\infty} \|f_n - f\|_{L^p} = 0 \quad \text{iff} \quad \lim_{n\to\infty} \|f_n\|_{L^p} = \|f\|_{L^p}.$$

For the forward implication, the reverse triangle inequality gives

$$\big| \|f_n\|_{L^p} - \|f\|_{L^p} \big| \le \|f_n - f\|_{L^p} \to 0$$

as $n \to \infty$.

Conversely, since $|f_n - f|^p \le 2^{p-1}(|f_n|^p + |f|^p)$ by convexity[2] of $x \mapsto |x|^p$, we may apply the reverse Fatou lemma to get

[2] Or, by (1.16),
$$|f_n - f|^p \le 2^p(|f_n|^p + |f|^p).$$

$$\limsup_{n\to\infty} \int_X |f_n - f|^p \, d\mu \le \int_X \limsup_{n\to\infty} |f_n - f|^p \, d\mu = 0,$$

so that $\lim_{n\to\infty} \|f_n - f\|_{L^p} = 0$ as needed.

Exercise 1.3.8. The argument is similar to the proof of Proposition 1.3.8, except that we exclude the step where horizontal truncation is used to limit our consideration to bounded L^∞ functions of finite measure support. This is because in L^∞, functions do not necessarily 'decay at infinity.' Such decay allows us to use Markov's inequality to write the support of any L^p function with $0 < p < \infty$ as a countable union of finite measure sets, which lets us use horizontal truncation for measure spaces (245A exercise 1.4.36(x)). Consider the measure space with a singleton set $\{*\}$, where $\mu(\{*\}) = \infty$. Then $1_{\{*\}} \in L^\infty$, and $\int_{\{*\}} 1_{\{*\}} \, d\mu = \infty$, but the only function with finite measure support is the zero function, which has integral equal to zero.

Exercise 1.3.9. [I learned the following answer from `https://math.stackexchange.com/a/538087/`. The key idea I missed was that one could use a generating set to approximate a σ-algebra.] Suppose $\mathcal{X} = \langle \mathcal{A} \rangle$ for some countable set $\mathcal{A}$. By σ-finiteness, we may partition

$X = \bigcup_{n=1}^{\infty} X_n$ with $\mu(X_n) < \infty$. We prove that $L^p(X_n, \mathcal{X}\downharpoonright_{X_n}, \mu\downharpoonright_{X_n})$ is separable. Let $\epsilon > 0$ and $f \in L^p(X_n)$. Choose a simple function $f' = \sum_{j=1}^m c_j 1_{E_j}$ with rational coefficients c_j such that $\|f - f'\|_{L^p(X_n)} \leq \epsilon/2$. We define $\mathcal{A}_n := \{A \cap X_n : A \in \mathcal{A}\}$. Then $\mathcal{A}_n$ is countable, with $\mathcal{X}\downharpoonright_{X_n} = \langle \mathcal{A}_n \rangle$. Thus, by 245A exercise 1.4.28, we may approximate E_j by some set $E_j' \in \mathcal{A}_n$, so that $\mu(E_j \triangle E_j') \leq (\epsilon/2m|c_j|)^p$. Defining $f'' := \sum_{j=1}^m c_j 1_{E_j'}$, we see that

$$\begin{aligned}
\|f - f''\|_{L^p(X_n)} &\leq \|f - f'\|_{L^p(X_n)} + \|f' - f''\|_{L^p(X_n)} \\
&\leq \epsilon/2 + \sum_{j=1}^m |c_j| \|1_{E_j} - 1_{E_j'}\|_{L^p(X_n)} \\
&\leq \epsilon/2 + \sum_{j=1}^m |c_j| \|1_{E_j \triangle E_j'}\|_{L^p(X_n)} \\
&\leq \epsilon/2 + \sum_{j=1}^m |c_j| (\epsilon/2m|c_j|) \\
&= \epsilon.
\end{aligned}$$

Since the set D_n of rational linear combinations $\sum_{j=1}^m c_j 1_{E_j}$ of indicator functions 1_{E_j} with sets $E_j \in \mathcal{A}_n$ is countable, it follows that $L^p(X_n)$ is separable.

The general case then follows from letting D be the set of finite sums of simple functions with at most one taken from each D_n; that is, we let $D = \bigcup_{N=1}^{\infty} \{\sum_{n=1}^N f_n : f_n \in D_n\}$. Then D is countable, and given $f \in L^p(X)$, we may choose an approximation of $f\downharpoonright_{X_n}$ by some function $f_n \in D_n$ such that $\|f\downharpoonright_{X_n} - f_n\|_{L^p(X_n)} \leq \epsilon/2^{n+1}$, so that $\sum_{n=1}^{\infty} \|f\downharpoonright_{X_n} - f_n\|_{L^p(X_n)} \leq \epsilon/2$. By the completeness of L^p, we then have

$$\sum_{n=1}^{\infty} (f\downharpoonright_{X_n} - f_n) = f - \sum_{n=1}^{\infty} f_n \in L^p(X),$$

and so $\sum_{n=1}^{\infty} f_n \in L^p(X)$ as well. Thus we may choose sufficiently large N for which $\sum_{n=1}^N f_n \in D$ is a good approximation for f, so that $\|f - \sum_{n=1}^N f_n\|_{L^p(X)} \leq \epsilon$ as needed.

We note that L^∞ need not be separable. Consider $(\mathbf{N}, 2^{\mathbf{N}}, \#)$ for example, where we have $\|f\|_{L^\infty} = \sup_{n \in \mathbf{N}} |f(n)|$. If we look at the uncountably many maps of the form $f \colon \mathbf{N} \to \{0, 1\} \subset \mathbf{C}$, we see that they all belong to L^∞. Given any two distinct maps f and g of this form, we see that $\|f - g\|_{L^\infty} = 1$. Thus we may take small open balls in L^∞ around each function of this form, to obtain uncountably many disjoint open sets. It follows that $L^\infty(\mathbf{N}, 2^{\mathbf{N}}, \#)$ is not separable.

Exercise 1.3.10. Let us first note that we are dealing with Young's inequality: if $a, b \geq 0$ are nonnegative real numbers and $p, q > 1$ are dual (so that $1/p + 1/q = 1$), then $ab \leq a^p/p + b^q/q$.

Consider the use of convexity in the proof of Hölder's inequaliy. In particular, we used the fact that

$$e^{(1-t)\alpha + t\beta} \leq (1-t)e^\alpha + te^\beta.$$

When $0 < t < 1$, equality holds iff $\alpha = \beta$. Since we used this with $\alpha = p \log |f(x)|$ and $\beta = q \log |g(x)|$, it follows that $|f(x)|^p = |g(x)|^q$.

Thus the claim follows from the normalizations of $|f|$ and $|g|$ as in the proof.

Alternatively, one could see this geometrically by proving a more general form of Young's inequality: given a real-valued continuous strictly increasing function $f\colon [0,a] \to [0,+\infty)$ with $f(0) = 0$, we have

$$ab \le \int_0^a f(x)\,dx + \int_0^b f^{-1}(x)\,dx,$$

where $b \in \operatorname{im} f$. Indeed, the areas given by the two integrals cover the rectangle $[0,a] \times [0,b]$, which gives the result geometrically. Equality then holds iff $b = f(a)$, and we may recover the equality case for Hölder's inequality by setting $f(x) = x^{p/q}$.

If p is infinite, then we get

$$\|fg\|_{L^q} \le \|f\|_{L^\infty}\|g\|_{L^q},$$

or

$$\Big(\int_X |f(x)g(x)|^q\,d\mu\Big)^{1/q} \le \|f\|_{L^\infty}\Big(\int_X |g(x)|^q\,d\mu\Big)^{1/q}.$$

Thus equality holds iff $|f| = \|f\|_{L^\infty}$; that is, if $|f|$ is constant a.e..

Exercise 1.3.11. For $q = p$ the result is clear, so suppose $0 < q < p$. By Hölder's inequality, we have

$$\|f\|_{L^q} \le \|1_E\|_{L^{pq/(p-q)}}\|f\|_{L^p} = \mu(E)^{1/q-1/p}\|f\|_{L^p}.$$

Equality holds iff $|f| = 1_E$.

Exercise 1.3.12. [This problem is hard! The solution below is from `https://math.stackexchange.com/a/669971/`.]

The idea is to use level sets. Let $E_\lambda := \{x \in X : |f(x)| \ge \lambda\}$, and suppose $0 < p < q < \infty$. Then

$$\|f\|_{L^p}^p = \int_X |f(x)|^p\,d\mu \ge \int_{E_\lambda} |f(x)|^p\,d\mu \ge \lambda^p \mu(E_\lambda).$$

In particular, if $\lambda > m^{-1/p}\|f\|_{L^p}$, then

$$\|f\|_{L^p}^p > m^{-1}\|f\|_{L^p}^p \mu(E_\lambda),$$

so that $\mu(E_\lambda) < m$. By definition of m, we must have $\mu(E_\lambda) = 0$. Thus

$$|f| \le m^{-1/p}\|f\|_{L^p} \quad \text{a.e..}$$

It follows that

$$\begin{aligned}\int_X |f(x)|^q\,d\mu &\le \||f|^{q-p}\|_{L^\infty}\int_X |f(x)|^p\,d\mu \\ &\le (m^{-1/p}\|f\|_{L^p})^{q-p}\int_X |f(x)|^p\,d\mu,\end{aligned}$$

so that

$$\begin{aligned}\|f\|_{L^q} &\le (m^{-1/p}\|f\|_{L^p})^{1-p/q}\Big(\int_X |f(x)|^p\,d\mu\Big)^{1/p+(1/q-1/p)} \\ &= m^{1/q-1/p}\|f\|_{L^p}\end{aligned}$$

as needed. Equality holds iff $|f|$ is constant. The case for $q = \infty$ then follows from taking the limit $q \to \infty$, noting that $\|f\|_{L^q} \le C < \infty$ for some constant C and sufficiently large q.

Exercise 1.3.13. By Hölder's inequality, we have

$$\begin{aligned} \|f\|_{L^p} &= \big\| |f|^{1-\theta} |f|^{\theta} \big\|_{L^p} \\ &\le \big\| |f|^{1-\theta} \big\|_{L^{p_0/(1-\theta)}} \big\| |f|^{\theta} \big\|_{L^{p_1/\theta}} \\ &= \|f\|_{L^{p_0}}^{1-\theta} \|f\|_{L^{p_1}}^{\theta}. \end{aligned}$$

Equality holds when $|f|^{p_0} = |f|^{p_1}$; that is, when $|f|^{p_1 - p_0} = 1$, or when $|f| = 1_X$.

Exercise 1.3.14. By exercise 1.3.11, $\|f\|_{L^p} \le \mu(E)^{1/p - 1/p_0} \|f\|_{L^{p_0}}$, so that

$$\|f\|_{L^p}^p \le \mu(E)^{1 - p/p_0} \|f\|_{L^{p_0}}^p.$$

Thus

$$\limsup_{p \to 0} \|f\|_{L^p}^p \le \mu(E).$$

By Fatou's lemma,

$$\liminf_{n \to \infty} \int_X |f(x)|^{1/n} \, d\mu \ge \int_X \liminf_{n \to \infty} |f(x)|^{1/n} \, d\mu = \mu(E),$$

and so $\liminf_{p \to 0} \|f\|_{L^p}^p = \liminf_{n \to \infty} \|f\|_{L^{1/n}}^{1/n} \ge \mu(E)$ by continuity, which gives the result.

Time is a waste of money.

— Oscar Wilde, *Phrases and Philosophies for the Use of the Young* (1894)

2.3. The Stone and Loomis–Sikorski Representation Theorems

When DEK taught Concrete Mathematics at Stanford for the first time, he explained the somewhat strange title by saying that it was his attempt to teach a math course that was hard instead of soft. He announced that, contrary to the expectations of some of his colleagues, he was not *going to teach the Theory of Aggregates, nor Stone's Embedding Theorem, nor even the Stone–Čech compactification. (Several students from the civil engineering department got up and quietly left the room.)*

— Ronald L. Graham, Donald E. Knuth, & Oren Patashnik, *Concrete Mathematics* (1988)

Exercise 2.3.1. Let X and Y be Stone spaces with isomorphic clopen algebras, with isomorphism $\phi\colon \mathrm{Cl}(X) \to \mathrm{Cl}(Y)$. Given $x \in X$, we define the set $F(x) := \bigcap_{x\in K\in\mathrm{Cl}(X)} \phi(K) \subset Y$. Since $\phi(K_1 \cap \cdots \cap K_n) = \phi(K_1) \cap \cdots \cap \phi(K_n)$ by definition of an abstract Boolean morphism, the finite intersection property implies that $F(x)$ is non-empty. Now suppose that $F(x)$ contained distinct points a and b. Then, normality gives us disjoint open sets $a \in U$ and $b \in V$. Since the clopen sets form a base for the topology on Y, there exist clopen sets $a \in K_a \subset U$ and $b \in K_b \subset V$; we may take their intersections with $F(x)$ to ensure they lie in $F(x)$. Then $\phi^{-1}(K_a), \phi^{-1}(K_b) \subset K$ for every $x \in K \in \mathrm{Cl}(X)$, and so

$$\phi^{-1}(K_a), \phi^{-1}(K_b) \subset \bigcap_{x\in K\in\mathrm{Cl}(X)} K = \{x\}.$$

Thus $K_a = K_b$, a contradiction. We conclude that $F(x)$ is a singleton, and we define $f\colon X \to Y$ by sending x to the single element of $F(x)$.

We may construct G and $g\colon Y \to X$ similarly, with $G(y) := \bigcap_{y\in K'\in\mathrm{Cl}(Y)} \phi^{-1}(K')$. Applying the above arguments, we see that $G(y)$ is a singleton as well, and so g is a well-defined function. We claim that f and g are inverses. To prove that $g \circ f = \mathrm{id}_X$, it suffices to prove that $x \in G(f(x))$, since $G(f(x))$ is a singleton.

Claim. *Given $K' \in \mathrm{Cl}(Y)$, we have $f^{-1}(K') \subset \phi^{-1}(K')$.*

Proof. Note that $\phi\colon \mathrm{Cl}(X) \to \mathrm{Cl}(Y)$ is an isomorphism that respects inclusion. We know that $\{f(x)\} = \bigcap_{x\in K\in\mathrm{Cl}(X)} \phi(K) \subset K'$. If we can find finitely many K_i from this collection such that

$$\phi(K_1) \cap \cdots \cap \phi(K_n) \subset K',$$

we would then have

$$\begin{aligned}\phi(K_1 \cap \cdots \cap K_n) \subset K' &\implies K_1 \cap \cdots \cap K_n \subset \phi^{-1}(K') \\ &\implies x \in K_1 \cap \cdots \cap K_n \subset \phi^{-1}(K')\end{aligned}$$

as needed. Consider the collection

$$(X \setminus K') \cap \bigcap_{x\in K\in\mathrm{Cl}(X)} \phi(K) = \varnothing$$

of closed sets. Then, by the finite intersection property, we have

$$(X \setminus K') \cap \phi(K_1) \cap \cdots \cap \phi(K_n) = \varnothing$$

for some sets $x \in K_i \in \mathrm{Cl}(X)$ with $1 \leq i \leq n$, as desired. □

We now know that $g \circ f = \mathrm{id}_X$, and we may argue similarly to prove that $g^{-1}(K) \subset \phi(K)$ for $K \in \mathrm{Cl}(X)$, so that $f \circ g = \mathrm{id}_Y$. Therefore $f \colon X \to Y$ is a bijection, and we will henceforth write f^{-1} instead of g. It remains to be shown that f is a homeomorphism. Since ϕ maps clopen sets to clopen sets, and since we may verify continuity of a map by checking that all preimages of basic open sets (from a base for the topology) are open, it suffices to prove the following:

Claim. *Given $K' \in \mathrm{Cl}(Y)$, we have $f^{-1}(K') = \phi^{-1}(K')$.*

Proof. We only need to prove that $\phi^{-1}(K') \subset f^{-1}(K')$. Suppose contrapositively that $f(x) \notin K'$. We prove that $x \notin \phi^{-1}(K')$. By normality, there exist disjoint open sets U and V such that $f(x) \in U$ and $K' \subset V$. By Lemma 2.3.3, clopen sets form a base for the topology on Y, and so $f(x) \in K \subset U$ for some clopen K. Thus $\phi^{-1}(K) \cap \phi^{-1}(K') = \varnothing$. Since $f(x) \in K$, we have

$$x = f^{-1}(f(x)) \in \bigcap_{f(x) \in K'' \in \mathrm{Cl}(Y)} \phi^{-1}(K'') \subset \phi^{-1}(K),$$

and so $x \notin \phi^{-1}(K')$, as needed. □

Thus, f is continuous. Applying the above argument to f^{-1}, we conclude that f is a homeomorphism as needed.

Exercise 2.3.2. This is the fact that any finite Boolean algebra is atomic, which can be seen by sending $b \in \mathcal{B}$ to the intersection of all sets of $\mathcal{B}$ containing b. [There is probably a way to get this result out of the Stone representation theorem, which I think involves proving that the Stone space in question is finite and thus discrete, but I haven't worked out the details.]

Exercise 2.3.3. [To do. . .]

Unfortunately, it appears that there is now in your world
a race of vampires, called referees, who clamp down mercilessly
upon mathematicians unless they know the right passwords.
I shall do my best to modernize my language and notations,
but I am well aware of my shortcomings in that respect;
I can assure you, at any rate, that my intentions are honourable
and my results invariant, probably canonical, perhaps even functorial.
But please allow me to assume that the characteristic is not 2.

— André Weil, in *Annals of Mathematics* **69** (1959)

1.4. Hilbert spaces

Dr. von Neumann,
I would very much like to know,
what after all is a Hilbert space?

— DAVID HILBERT, apocryphally (1929)

Exercise 1.4.1. In the real case, we have the identity

$$\langle x, y\rangle = \frac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2\big).$$

Thus, using linearity, we have

$$\langle T(x), T(y)\rangle = \frac{1}{4}\big(\|T(x+y)\|^2 - \|T(x-y)\|^2\big).$$

If the inner product is preserved, then by setting $x = y = v/2$, we have

$$\|T(v)\|^2 - \|T(0)\|^2 = \|v\|^2 - \|0\|^2,$$

and so the norm is preserved. Conversely, we may use the fact that $\|T(x+y)\| = \|x+y\|$ and $\|T(x-y)\| = \|x-y\|$.

The complex case is similar; we just use the identity

$$\langle x, y\rangle = \frac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2\big) + \frac{1}{4}\big(\|x+iy\|^2 - \|x-iy\|^2\big).$$

Exercise 1.4.2. Let $G = (\langle x_i, x_j\rangle)_{1\le i,j\le n}$ be the Gram matrix for $\langle\cdot,\cdot\rangle$. Then G is Hermitian, since $G_{ij} = \langle x_i, x_j\rangle = \overline{\langle x_j, x_i\rangle} = \overline{G_{ji}}$. To prove positive semi-definiteness, we compute

When I wrote this solution, I thought that $x^*Ay = \langle x, y\rangle$. But I think now that it's supposed to be $x^*Ay = \langle y, x\rangle$; that will probably give a solution with less conjugates.

$$z^*Az = \left\langle \sum_{i=1}^n \overline{z_i}x_i, \sum_{i=1}^n \overline{z_i}x_i \right\rangle \ge 0.$$

Suppose now that the vectors x_i are linearly dependent. Then we may write $z_1x_1 + \cdots + z_nx_n = 0$ for some $z = (z_1, \dots, z_n) \in \mathbf{C}^n \setminus \{0\}$, so that $\overline{z}^*G\overline{z} = \langle \sum_{i=1}^n z_ix_i, \sum_{i=1}^n z_ix_i\rangle = 0$. Conversely, if $z^*Gz = 0$ for some non-zero $z \in \mathbf{C}^n$, then the non-degeneracy of $\langle\cdot,\cdot\rangle$ implies that we have the non-trivial linear combination $\sum_{i=1}^n \overline{z_i}x_i = 0$.

[Had to look up the following to figure it out.] Let M be an $n\times n$ Hermitian matrix. Then M is positive semi-definite iff there exists a decomposition $M = B^*B$. If $M = B^*B$, then

The matrix B is invertible iff M is positive-definite.

$$z^*Mz = z^*B^*Bz = (Bz)^*(Bz) = \|Bz\|^2 \ge 0.$$

Conversely, since M is Hermitian, we may decompose $M = Q^{-1}DQ$ where Q is unitary and D is a diagonal matrix whose entries are the eigenvalues of M. By positive semi-definiteness, these eigenvalues are non-negative, and so we may define $D^{1/2}$. Setting $B := D^{1/2}Q$, we obtain

$$B^*B = Q^*(D^{1/2})^*D^{1/2}Q = Q^*DQ = Q^{-1}DQ = M$$

as needed.

We may then use the columns of B as our vectors x_i, since $\langle B_i, B_j\rangle = M_{ij}$ as needed.

Exercise 1.4.3. We compute

$$\begin{aligned}\|x_1 + x_2\|^2 &= \langle x_1 + x_2, x_1 + x_2\rangle \\ &= \langle x_1, x_1\rangle + \langle x_1, x_2\rangle + \langle x_2, x_1\rangle + \langle x_2, x_2\rangle \\ &= \|x_1\|^2 + \|x_2\|^2\end{aligned}$$

in the $n = 2$ case. Inductively, we then have

$$\begin{aligned}\Big\|\sum_{i=1}^n x_i\Big\|^2 &= \Big\langle \sum_{i=1}^n x_i, \sum_{i=1}^n x_i\Big\rangle \\ &= \Big\langle \sum_{i=1}^{n-1} x_i, \sum_{i=1}^{n-1} x_i\Big\rangle + \Big\langle \sum_{i=1}^{n-1} x_i, x_n\Big\rangle + \Big\langle x_n, \sum_{i=1}^{n-1} x_i\Big\rangle + \langle x_n, x_n\rangle \\ &= \Big\|\sum_{i=1}^{n-1} x_i\Big\|^2 + \sum_{i=1}^{n-1}\langle x_i, x_n\rangle + \sum_{i=1}^{n-1}\langle x_n, x_i\rangle + \|x_n\|^2 \\ &= \sum_{i=1}^{n-1} \|x_i\|^2.\end{aligned}$$

In particular, since $\|x_1 + x_2\|^2 = \|x_1\|^2 + \|x_2\|^2 \geq \|x_1\|^2$, we see that $\|x_1 + x_2\| \geq \|x_1\|$ whenever x_1 and x_2 are orthogonal.

Exercise 1.4.4. Suppose $\sum_{k=1}^n c_k e_{\alpha_k} = 0$. Then

$$0 = \Big\langle \sum_{k=1}^n c_k e_{\alpha_k}, e_{\alpha_j}\Big\rangle = \sum_{k=1}^n c_k \langle e_{\alpha_k}, e_{\alpha_j}\rangle = c_j$$

for $1 \leq j \leq n$, and so $(e_\alpha)_{\alpha\in A}$ is linearly independent.

Suppose $x = \sum_{k=1}^n c_k e_{\alpha_k}$. Then $c_j = \langle x, e_{\alpha_j}\rangle$ as noted earlier, and if $e_\alpha \neq e_{\alpha_j}$ for $1 \leq j \leq n$, then $\langle x, e_\alpha\rangle = \sum_{k=1}^n c_k \langle e_{\alpha_j}, e_\alpha\rangle = 0$. Thus

$$\sum_{\alpha\in A}\langle x, e_\alpha\rangle e_\alpha = \sum_{k=1}^n \langle x, e_{\alpha_k}\rangle e_{\alpha_k} = \sum_{k=1}^n c_k e_{\alpha_k} = x$$

as needed.

Now we may compute (with no worries as only finitely many terms are non-zero)

$$\begin{aligned}\|x\|^2 &= \langle x, x\rangle \\ &= \Big\langle \sum_{\alpha\in A}\langle x, e_\alpha\rangle e_\alpha, \sum_{\beta\in A}\langle x, e_\beta\rangle e_\beta\Big\rangle \\ &= \sum_{\alpha\in A}\sum_{\beta\in A}\langle x, e_\alpha\rangle\overline{\langle x, e_\beta\rangle}\langle e_\alpha, e_\beta\rangle \\ &= \sum_{\alpha\in A}|\langle x, e_\alpha\rangle|^2.\end{aligned}$$

Exercise 1.4.5. The correct idea is the natural one, where we subtract from v its components in e_i and then normalize it. That is, we define

$$e_{n+1} := \frac{v - \sum_{i=1}^n \langle v, e_i\rangle e_i}{\|v - \sum_{i=1}^n \langle v, e_i\rangle e_i\|}.$$

Here, the denominator is non-zero as v does not lie in the span of $e_1, \dots, e_n$. Then, clearly $\|e_{n+1}\| = 1$, and for $1 \le j \le n$ we have

$$\begin{aligned}\langle e_{n+1}, e_j\rangle &= \frac{1}{\|v - \sum_{i=1}^n \langle v, e_i\rangle e_i\|}\Big\langle v - \sum_{i=1}^n \langle v, e_i\rangle e_i, e_j\Big\rangle \\ &= \frac{1}{\|v - \sum_{i=1}^n \langle v, e_i\rangle e_i\|}\Big(\langle v, e_j\rangle - \sum_{i=1}^n \langle v, e_i\rangle\langle e_i, e_j\rangle\Big) \\ &= 0,\end{aligned}$$

so that $\{e_1, \dots, e_{n+1}\}$ is orthonormal, with span equal to $\{e_1, \dots, e_n, v\}$.

Therefore, given some n-dimensional complex inner product space V, we may take a basis $\{e_1', \dots, e_n'\}$ for V, and modify it so that it is orthonormal. Then we may define a map $V \to \mathbf{C}^n$ by sending e_i' to $e_i \in \mathbf{C}^n$; this is seen to be invertible. Since this mapping preserves the inner product ($\langle e_i', e_j'\rangle = [i = j] = \langle e_i, e_j\rangle$), we conclude that V is isomorphic to $\mathbf{C}^n$ as a complex inner product space.

Exercise 1.4.6. We have

$$\begin{aligned}\|x+y\|^2 + \|x-y\|^2 &= \langle x+y, x+y\rangle + \langle x-y, x-y\rangle \\ &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &\quad + \langle x, x\rangle - \langle x, y\rangle - \langle y, x\rangle + \langle y, y\rangle \\ &= 2\|x\|^2 + 2\|y\|^2\end{aligned}$$

on any inner product space.

Suppose A and B are disjoint sets of finite positive measure in $L^p(X, \mathcal{X}, \mu)$ with $p \neq 2$. Then

$$\Big(\int_X |1_A + 1_B|^p \, d\mu\Big)^{1/p} + \Big(\int_X |1_A - 1_B|^p \, d\mu\Big)^{1/p} = 2(\mu(A) + \mu(B))^{1/p},$$

whereas

$$2\Big(\int_X |1_A|^p \, d\mu\Big)^{1/p} + 2\Big(\int_X |1_B|^p \, d\mu\Big)^{1/p} = 2(\mu(A)^{1/p} + \mu(B)^{1/p}).$$

[For the proof of the Hanner inequalities, I have nothing to say, so I refer the reader to Lieb–Loss section 2.5.]

Exercise 1.4.7. Suppose $S \subset H$ is a subspace that is also a Hilbert space. Then, given a convergent sequence $(x_n)_{n=1}^\infty \subset S$ with limit $x \in H$, we see that $(x_n)_{n=1}^\infty$ is a Cauchy sequence that lies completely in S. Thus it converges to a limit $x' \in S$, and uniqueness of limits implies that $x = x' \in S$ as needed. Conversely, suppose we are given a Cauchy sequence $(x_n)_{n=1}^\infty \subset S$. Then it is also Cauchy in H, and thus converges to a limit $x \in H$. Since S is a closed subset of H, it follows that $x \in S$, and so $(x_n)_{n=1}^\infty$ converges to a limit in S, as needed.

In particular, if $D \subset H$ is a closed dense subset, then $D = \overline{D} = H$. Thus, proper dense subspaces of Hilbert spaces are not Hilbert spaces.

Exercise 1.4.8. (Sketch) The following construction is very much like the construction of the real numbers as the *Cauchy completion* of the rationals, or more generally the *metric completion* of a metric space. Elements of $\overline{V}$ will be equivalence classes of Cauchy sequences in V; vectors $f \in V$ will correspond to the constant sequence $(f)_{n=1}^{\infty}$. Write $\|\cdot\|$ for the norm induced by the inner product $\langle \cdot, \cdot \rangle$. Given Cauchy sequences $(f_n)_{n=1}^{\infty}$ and $(g_n)_{n=1}^{\infty}$, we write $(f_n)_{n=1}^{\infty} \sim (g_n)_{n=1}^{\infty}$ if we have $\lim_{n\to\infty} \|f_n - g_n\| = 0$. This is an equivalence relation by the triangle inequality; we define $\overline{V}$ to be the space of Cauchy sequences in V quotiented by this relation. Elements of $\overline{V}$ are written as $[(f_n)]$. We define addition and scalar multiplication as expected, with $[(f_n)] + [(g_n)] := [(f_n + g_n)]$ and $c[(f_n)] := [(cf_n)]$. These operations are easily checked to be well-defined. We define the inner product

$$\langle [(f_n)], [(g_n)] \rangle := \lim_{n\to\infty} \langle f_n, g_n \rangle.$$

Clearly this extends the inner product on V (with elements f of V identified with constant sequences $[(f)]$). The limit exists, since we may verify that $(\langle f_n, g_n \rangle)_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbf{C}$. The idea is to compute

$$\begin{aligned} |\langle f_m, g_m \rangle - \langle f_n, g_n \rangle| &\leq |\langle f_m, g_m \rangle - \langle f_m, g_n \rangle| + |\langle f_m, g_n \rangle - \langle f_n, g_n \rangle| \\ &= |\langle f_m, g_m - g_n \rangle| + |\langle f_m - f_n, g_n \rangle| \\ &\leq \|f_m\| \|g_m - g_n\| + \|f_m - f_n\| \|g_n\|; \end{aligned}$$

this quantity can be made arbitrarily small for $m, n \geq N$ as Cauchy sequences are bounded. We may also verify that the inner product is well-defined. The inner product axioms follow easily from those of the original inner product on V. It remains to be proven that $\overline{V}$ is complete. Note that we have

$$\langle [(f_n)], [(f_n)] \rangle = \lim_{n\to\infty} \|f_n\|^2;$$

in particular, the norm on $\overline{V}$ is given by

$$\|[(f_n)]\| = \lim_{n\to\infty} \|f_n\|.$$

Let $([(f_{n,k})_{k=1}^{\infty}])_{n=1}^{\infty}$ be a Cauchy sequence in $\overline{V}$. Then, given $\epsilon > 0$, there exists N such that

$$\lim_{k\to\infty} \|f_{m,k} - f_{l,k}\| \leq \epsilon \tag{$*$}$$

whenever $m, l \geq N$. We construct a candidate limit $[(f_n)_{n=1}^{\infty}]$ for this sequence. Let $f_1 := f_{1,1}$. Since $(f_{2,k})_{k=1}^{\infty}$ is Cauchy, choose $n_2 > 1$ such that $\|f_{2,m} - f_{2,l}\| \leq 1/2$ whenever $m, l \geq n_2$. We then let $f_2 := f_{2,n_2}$. We may continue in this fashion to obtain a sequence $1 =: n_1 < n_2 < \cdots$ of positive integers with $\|f_{k,m} - f_{k,l}\| \leq 1/k$ whenever $m, l \geq n_k$ and $f_k := f_{k,n_k}$. The sequence $(f_n)_{n=1}^{\infty}$ is Cauchy since, for sufficiently large l and m with $l \geq m$, we have by $(*)$:

$$\begin{aligned} \|f_m - f_l\| &= \|f_{m,n_m} - f_{l,n_l}\| \\ &\leq \|f_{m,n_m} - f_{m,n_l}\| + \|f_{m,n_l} - f_{l,n_l}\| \\ &\leq \frac{1}{m} + \epsilon. \end{aligned}$$

Finally, we prove that $([(f_{n,k})_{k=1}^\infty])_{n=1}^\infty$ converges to $[(f_k)_{k=1}^\infty]$. We must find large N for which $\lim_{k\to\infty} \|f_{n,k} - f_k\| \le \epsilon$ whenever $n \ge N$. By $(*)$, we may choose N such that $\lim_{k\to\infty} \|f_{m,k} - f_{l,k}\| \le \epsilon/2$ for $m, l \ge N$, with $1/N < \epsilon/2$. Then, for $n \ge N$, we have

$$\begin{aligned}\lim_{k\to\infty} \|f_{n,k} - f_{k,n_k}\| &\le \lim_{k\to\infty} \|f_{n,k} - f_{n,n_k}\| + \lim_{k\to\infty} \|f_{n,n_k} - f_{k,n_k}\| \\ &\le \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\end{aligned}$$

as needed. Therefore $\overline{V}$ is complete.

We prove that V is dense in $\overline{V}$. Given $[(f_n)] \in \overline{V}$ and $\epsilon > 0$, we have $\|f_m - f_n\| \le \epsilon$ for $m, n \ge N$, so that the constant sequence $[(f_N)_{n=1}^\infty]$ is within ϵ of $[(f_n)]$.

Suppose $\overline{V}'$ is another completion of V, in the sense that $\overline{V}'$ contains a dense subspace isomorphic to V. Then we may define a map $\phi\colon \overline{V} \to \overline{V}'$ by sending $[(f_n)]$ to the limit of (f_n) in $\overline{V}'$; clearly this fixes V. This map is well-defined since if we write $f = \phi([(f_n)])$, and if $(f_n') \sim (f_n)$, then $\lim_{n\to\infty} \|f_n' - f\| = \lim_{n\to\infty} \|f_n' - f_n\| + \lim_{n\to\infty} \|f_n - f\| = \lim_{n\to\infty} \|f_n - f\|$. Since V is dense in $\overline{V}'$, given $f \in \overline{V}'$, we may choose for each n some $f_n \in V$ such that $\|f_n - f\|_{V'} \le 1/n$; this defines a map $\psi\colon \overline{V}' \to \overline{V}$ with $\psi\colon f \mapsto [(f_n)]$. This map can be checked to be well-defined. We can verify that $\phi \circ \psi = \mathrm{id}_{\overline{V}'}$ and $\psi \circ \phi = \mathrm{id}_{\overline{V}}$.

I forgot to prove that the inner product is preserved; this should be a standard continuity argument.

Exercise 1.4.9.

- Positivity of $\langle \cdot, \cdot \rangle_{H \oplus H'}$:

$$\langle (x, x'), (x, x') \rangle_{H\oplus H'} := \langle x, x \rangle_H + \langle x', x' \rangle_{H'} \ge 0.$$

- Sesquilinearity of $\langle \cdot, \cdot \rangle_{H \oplus H'}$:

We have

$$\begin{aligned}&\langle c(x, x') + d(y, y'), (z, z') \rangle_{H\oplus H'} \\ &\quad = \langle (cx + dy, cx' + dy'), (z, z') \rangle_{H\oplus H'} \\ &\quad = \langle cx + dy, z \rangle_H + \langle cx' + dy', z' \rangle_{H'} \\ &\quad = c\langle x, z \rangle_H + d\langle y, z \rangle_H + c\langle x', z' \rangle_{H'} + d\langle y', z' \rangle_{H'} \\ &\quad = c\langle (x, x'), (z, z') \rangle_{H\oplus H'} + d\langle (y, y'), (z, z') \rangle_{H\oplus H'};\end{aligned}$$

the conjugate linearity of the second slot is proven similarly.

- Conjugate symmetry of $\langle \cdot, \cdot \rangle_{H \oplus H'}$:

$$\begin{aligned}\langle (x, x'), (y, y') \rangle_{H\oplus H'} &:= \langle x, y \rangle_H + \langle x', y' \rangle_{H'} \\ &= \overline{\langle y, x \rangle_H} + \overline{\langle y', x' \rangle_{H'}} \\ &= \overline{\langle (y, y'), (x, x') \rangle_{H\oplus H'}}\end{aligned}$$

- Completeness of $H \oplus H'$:

The norm in $H \oplus H'$ is given by

$$\|(x, x')\| = \sqrt{\langle (x, x'), (x, x') \rangle} = \sqrt{\|x\|^2 + \|x'\|^2}.$$

Let $((x_n, x'_n))_{n=1}^\infty$ be a Cauchy sequence. Then, for every $\epsilon > 0$, there exists N such that

$$\|(x_m - x_n, x'_m - x'_n)\| = \sqrt{\|x_m - x_n\|^2 + \|x'_m - x'_n\|^2} \leq \epsilon$$

whenever $m, n \geq N$. In particular, we have $\|x_m - x_n\| \leq \epsilon$ and $\|x'_m - x'_n\| \leq \epsilon$ whenever $m, n \geq N$, so that $(x_n)_{n\in\mathbf{N}}$ and $(x'_n)_{n\in\mathbf{N}}$ are Cauchy. Thus they converge to limits $x \in H$ and $x' \in H'$, and it is easy to show that $(x_n, x'_n) \to (x, x')$ as needed.

Exercise 1.4.10.

- K is convex but not closed.

Let $H = \mathbf{R}$, $K = (0, 1)$, and $x = 0$. Then $d(x, K) = 0$, but all points of K are at a positive distance from x.

- K is closed but not convex.

[Had to look this up.] Let $H = \ell^2(\mathbf{N})$, $K = \{(1 + 1/n)e_n : n \in \mathbf{N}\}$, and $x = (0)_{n\in\mathbf{N}}$. Then

$$d(x, K) = \inf_{n\in\mathbf{N}} \|(1 + 1/n)e_n\| = \inf_{n\in\mathbf{N}} (1 + 1/n) = 1.$$

Since $d((1 + 1/n)e_n, (1 + 1/m)e_m) \geq \sqrt{2}$ for distinct points of K, we see that K consists solely of isolated points, and so K is closed.

- K is closed convex, but H is not complete.

Let $H = C([0, 1]) \subset L^2([0, 1])$, and let K be the subspace of continuous functions supported on $[0, 1/2]$. Then ...

- Existence (but not uniqueness) can be recovered if K is assumed to be compact rather than convex.

Let $D := \inf_{y\in K} \|x - y\|$ as in the original proof, and find a sequence $y_n \in K$ such that $\|x - y_n\| \to D$. Use compactness to extract a convergent subsequence $y_{n_j} \to y$. Then $y \in K$ since K is closed, and $\|x - y\| = D$.

Exercise 1.4.11. [To do...]

Exercise 1.4.12. The subspace V is convex by linearity, and so given $x \in H$ there exists a minimizer $x_V \in V$. Clearly x_V is the closest element of V to x. Let $x_{V^\perp} := x - x_V$. Suppose for contradiction that $\langle x_{V^\perp}, v\rangle \neq 0$ for some $v \in V$. Scale v so that $\|v\| = 1$, and set

$$\begin{cases} x'_{V^\perp} := x_{V^\perp} - \langle x_{V^\perp}, v\rangle v, \\ x'_V := x_V + \langle x_{V^\perp}, x\rangle v. \end{cases}$$

Then $x = x'_V + x'_{V^\perp}$ with $x'_V \in V$. We have

$$\langle x'_{V^\perp}, v\rangle = \langle x_{V^\perp} - \langle x_{V^\perp}, v\rangle v, v\rangle = \langle x_{V^\perp}, v\rangle - \langle x_{V^\perp}, v\rangle\langle v, v\rangle = 0,$$

and so $x'_{V^\perp}$ is orthogonal to $\langle x_{V^\perp}, v\rangle v$. The Pythagorean theorem then implies that

$$\|x_{V^\perp}\|^2 = \|x'_{V^\perp}\|^2 + \|\langle x_{V^\perp}, v\rangle v\|^2,$$

so that $\|x'_{V^\perp}\| < \|x_{V^\perp}\|$, a contradiction. Thus $x_{V^\perp}$ is orthogonal to every element of V as needed.

Exercise 1.4.13. Let V be a subspace of a Hilbert space H.

- $V^\perp$ is a closed subspace of H, and $(V^\perp)^\perp$ is the closure of V.

By sesquilinearity, $V^\perp$ is a subspace of H. By continuity of the inner product, $V^\perp$ is closed. Thus $(V^\perp)^\perp$ is a closed subspace. It contains V since, if $v \in V$ and $x \in V^\perp$, then $\langle v, x\rangle = 0$. Thus $\overline{V} \subset (V^\perp)^\perp$. Conversely, if $x \in (V^\perp)^\perp$, then we may write $x = x_{\overline{V}} + x_{\overline{V}^\perp}$, with

$$\langle x, x_{\overline{V}^\perp}\rangle = \langle x_{\overline{V}}, x_{\overline{V}^\perp}\rangle + \langle x_{\overline{V}^\perp}, x_{\overline{V}^\perp}\rangle.$$

Since $\overline{V}^\perp = V^\perp$, it follows that $\|x_{\overline{V}^\perp}\| = 0$, and so $x = x_{\overline{V}} \in \overline{V}$ as needed.

- $V^\perp$ is the trivial subspace $\{0\}$ if and only if V is dense.

If $V^\perp = \{0\}$, then $(V^\perp)^\perp = H$. Thus $\overline{V} = H$, and so V is dense. Conversely, let $w \in V^\perp$, and use the density of V to choose a sequence $(w_n)_{n=1}^\infty$ in V converging to w. Then, we have $0 = \langle w_n, w\rangle \to \langle w, w\rangle$, and so $w = 0$ by continuity.

- If V is closed, then H is isomorphic to the direct sum of V and $V^\perp$.

The obvious candidate for the isomorphism is the map $\phi\colon H \to V \oplus V^\perp$ defined by $\phi(x) := (x_V, x_{V^\perp})$, which is well-defined as V is closed. It has an inverse given by $(v, w) \mapsto v + w$, and the inner product is preserved since

$$\begin{aligned}\langle \phi(x), \phi(y)\rangle &= \langle (x_V, x_{V^\perp}), (y_V, y_{V^\perp})\rangle \\ &= \langle x_V, y_V\rangle + \langle x_{V^\perp}, y_{V^\perp}\rangle \\ &= \langle x_V, y_V\rangle + \langle x_V, y_{V^\perp}\rangle + \langle x_{V^\perp}, y_V\rangle + \langle x_{V^\perp}, y_{V^\perp}\rangle \\ &= \langle x, y\rangle.\end{aligned}$$

- If V, W are two closed subspaces of H, then $(V+W)^\perp = V^\perp \cap W^\perp$ and $(V \cap W)^\perp = \overline{V^\perp + W^\perp}$.

If $x \in (V+W)^\perp$, then $\langle x, v+w\rangle = 0$ for all $v \in V$ and $w \in W$. In particular, since $0 \in V$ and $0 \in W$, we have $\langle x, v\rangle = \langle x, w\rangle = 0$ for all $v \in V$ and $w \in W$. Thus $x \in V^\perp \cap W^\perp$. Conversely, since $0 = \langle x, v\rangle + \langle x, w\rangle = \langle x, v+w\rangle$, we have $V^\perp \cap W^\perp \subset (V+W)^\perp$.

If $x \in (V \cap W)^\perp$, then $\langle x, y\rangle = 0$ whenever $y \in V \cap W$. Writing $x = x_V + x_{V^\perp}, \ldots$ [To do!]

Conversely, if $x \in V^\perp$ and $y \in W^\perp$,

$$\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle = 0$$

whenever $z \in V \cap W$. Thus $V^\perp + W^\perp \subset (V \cap W)^\perp$. By continuity of the inner product, we deduce that $\overline{V^\perp + W^\perp} \subset (V \cap W)^\perp$.

Exercise 1.4.14. [To do...]

Remark. The following fact will be quite useful for the next few exercises: if $\langle x, v\rangle = \langle x', v\rangle$ for all $v \in H$, then $x = x'$. Indeed, setting $v := x - x'$, we see that $\langle x - x', x - x'\rangle = 0$. Non-degeneracy then implies that $x - x' = 0$, so that $x = x'$ as needed.

Exercise 1.4.15. We define $T^\dagger \colon H' \to H$ as follows: given $y \in H'$, the map $\lambda \colon H \to \mathbf{C}$ defined by $\lambda(x) = \langle T(x), y\rangle$ is a continuous linear functional. The Riesz representation theorem gives us a unique element $z \in H$ such that $\lambda = \lambda_z$. Then we define $T^\dagger(y) := z$. Thus $\langle T(x), y\rangle = \langle x, T^\dagger(y)\rangle$ for all $x \in H$ and $y \in H'$. Additivity follows from the fact that

$$\begin{aligned}\langle x, T^\dagger(y+y')\rangle &= \langle T(x), y+y'\rangle \\ &= \langle T(x), y\rangle + \langle T(x), y'\rangle \\ &= \langle x, T^\dagger(y)\rangle + \langle x, T^\dagger(y')\rangle \\ &= \langle x, T^\dagger(y) + T^\dagger(y')\rangle;\end{aligned}$$

scalar multiplication is verified similarly. Thus $T^\dagger$ is linear. To verify continuity, suppose $y_n \to y$. Then the continuity of the inner product implies that

$$\begin{aligned}\langle x, \lim_{n\to\infty} T^\dagger(y_n) - T^\dagger(y)\rangle &= \lim_{n\to\infty} \langle x, T^\dagger(y_n)\rangle - \langle x, T^\dagger(y)\rangle \\ &= \lim_{n\to\infty} \langle T(x), y_n - y\rangle \\ &= \langle T(x), 0\rangle \\ &= 0,\end{aligned}$$

and so $\lim_{n\to\infty} T^\dagger(y_n) = T^\dagger(y)$.

Exercise 1.4.16.

- $(T^\dagger)^\dagger = T$.

We have

$$\langle x, (T^\dagger)^\dagger(x')\rangle = \langle T^\dagger(x), x'\rangle = \overline{\langle x', T^\dagger(x)\rangle} = \overline{\langle T(x'), x\rangle} = \langle x, T(x')\rangle$$

for all $x, x' \in H$.

- T is an isometry iff $T^\dagger T = \mathrm{id}_H$.

Suppose T is an isometry. Then

$$\langle x, T^\dagger T(x')\rangle = \langle T(x), T(x')\rangle = \langle x, x'\rangle = \langle x, \mathrm{id}_H(x')\rangle.$$

Conversely, if $T^\dagger T = \mathrm{id}_H$, then

$$\langle T(x), T(x')\rangle = \langle x, T^\dagger T(x')\rangle = \langle x, x'\rangle$$

as needed.

- T is an isomorphism iff $T^\dagger T = \mathrm{id}_H$ and $TT^\dagger = \mathrm{id}_{H'}$.

Suppose that T is an isomorphism. Since T preserves the inner product, we have

$$\langle x, T^{-1}(y)\rangle = \langle T(x), y\rangle = \langle x, T^\dagger y\rangle.$$

Thus $T^{-1} = T^\dagger$, and so $T^\dagger T = \mathrm{id}_H$ and $TT^\dagger = \mathrm{id}_{H'}$. Conversely, we see that the inverse of T exists, with $T^{-1} = T^\dagger$, and so

$$\langle T(x), T(x')\rangle = \langle x, T^\dagger T(x')\rangle = \langle x, x'\rangle.$$

Thus T is an invertible isometry; that is, an isomorphism.

- If $S\colon H' \to H''$ is a continuous linear transformation, then $(ST)^\dagger = T^\dagger S^\dagger$.

We compute

$$\langle x, (ST)^\dagger(z)\rangle = \langle S(T(x)), z\rangle = \langle T(x), S^\dagger(z)\rangle = \langle x, T^\dagger(S^\dagger(z))\rangle.$$

Exercise 1.4.17. Write $x = \pi_V(x) + x_{V^\perp}$. Then

$$\langle \pi_V(x), v\rangle = \langle x - x_{V^\perp}, v\rangle = \langle x, v\rangle = \langle x, \iota_V(v)\rangle.$$

Exercise 1.4.18. (i) Suppose $\sum_{n=1}^\infty |c_n|^2 < \infty$. Then, for large N, we have $\sum_{n=N}^\infty |c_n|^2 \le \epsilon$. Thus, the Pythagorean theorem implies that

$$\Big\|\sum_{n=k}^{l} c_n e_n\Big\|^2 = \sum_{n=k}^{l} \|c_n e_n\|^2 = \sum_{n=k}^{l} |c_n|^2 \le \epsilon$$

whenever $k, l \ge N$ as needed. Therefore completeness implies that $\sum_{n=1}^\infty c_n e_n$ exists.

Conversely, suppose $\sum_{n=1}^\infty c_n e_n$ exists. Then, we may compute

$$\begin{aligned}\sum_{n=1}^\infty |c_n|^2 &= \sum_{n=1}^\infty \|c_n e_n\|^2 \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \|c_n e_n\|^2 \\ &= \lim_{N\to\infty} \Big\|\sum_{n=1}^{N} c_n e_n\Big\|^2 \\ &= \Big\|\sum_{n=1}^{\infty} c_n e_n\Big\|^2 \\ &< \infty.\end{aligned}$$

(ii) Since $\sum_{n=1}^\infty |c_n|^2$ is absolutely convergent in $\mathbf{R}$, it is conditionally convergent as well, so that $\sum_{n=1}^\infty |c_{\sigma(n)}|^2 = \sum_{n=1}^\infty |c_n|^2 < \infty$ for any permutation $\sigma\colon \mathbf{N} \to \mathbf{N}$.

We now prove that $\lim_{N\to\infty} \sum_{n=1}^{N} c_{\sigma(n)} e_{\sigma(n)} = \sum_{n=1}^\infty c_n e_n$. Let $\epsilon > 0$, and choose large M such that

$$\Big\|\sum_{n=M+1}^{\infty} c_n e_n\Big\| \le \epsilon/2$$

and
$$\sum_{n=M+1}^{\infty} |c_n|^2 \le (\epsilon/2)^2.$$
Then, choose $N > M$ such that
$$\{1, \dots, M\} \subset \{\sigma(1), \dots, \sigma(N)\}.$$
Writing $S := \{1, \dots, N\} \setminus \{k : 1 \le \sigma(k) \le M\}$, it follows that
$$\begin{aligned}
\Big\| \sum_{n=1}^{\infty} c_n e_n - \sum_{n=1}^{N} c_{\sigma(n)} e_{\sigma(n)} \Big\| &= \Big\| \sum_{n=M+1}^{\infty} c_n e_n - \sum_{n\in S} c_{\sigma(n)} e_{\sigma(n)} \Big\| \\
&\le \Big\| \sum_{n=M+1}^{\infty} c_n e_n \Big\| + \Big\| \sum_{n\in S} c_{\sigma(n)} e_{\sigma(n)} \Big\| \\
&\le \epsilon/2 + \Big(\sum_{n\in S} |c_{\sigma(n)}|^2 \Big)^{1/2} \\
&\le \epsilon,
\end{aligned}$$
as needed.

(iii) Define $\phi\colon \ell^2(\mathbf{N}) \to H$ by $\phi\colon (c_n)_{n=1}^{\infty} \mapsto \sum_{n=1}^{\infty} c_n e_n$; this is well-defined by (i). Then, by continuity of the inner product, we have
$$\begin{aligned}
\langle \phi((c_m)_{m=1}^{\infty}), \phi((c'_n)_{n=1}^{\infty}) \rangle &= \Big\langle \sum_{m=1}^{\infty} c_m e_m, \sum_{n=1}^{\infty} c'_n e_n \Big\rangle \\
&= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_m c'_n \langle e_m, e_n \rangle \\
&= \sum_{n=1}^{\infty} c_n \overline{c'_n} \\
&= \langle (c_m)_{m=1}^{\infty}, (c'_n)_{n=1}^{\infty} \rangle.
\end{aligned}$$

(iv) If $x \in V$, then $x = \sum_{n=1}^{\infty} c_n e_n$. Since $\langle x, e_n \rangle = \langle \sum_{i=1}^{\infty} c_i e_i, e_n \rangle = \sum_{i=1}^{\infty} c_i \langle e_i, e_n \rangle = c_n$, it follows that $x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$. If $x \in H$, then $\pi_V(x) \in V$, and so by exercise 1.4.17, we have
$$\pi_V(x) = \sum_{n=1}^{\infty} \langle \pi_V(x), e_n \rangle e_n = \sum_{n=1}^{\infty} \langle x, \iota_V(e_n) \rangle e_n = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$$
as needed. We compute
$$\|\pi_V(x)\| = \Big(\sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 \Big)^{1/2}.$$
We also have by the Pythagorean theorem
$$\sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 = \|\pi_V(x)\|^2 \le \|x\|^2.$$

Exercise 1.4.19. We first prove that (i) is equivalent to (ii). Given $x \in H$, we have $x = \sum_{\alpha\in A} c_\alpha e_\alpha$, which we may rewrite as $\sum_{n=1}^{\infty} c_{\alpha_n} e_{\alpha_n}$ since at most countably many terms are non-zero. This is the limit of finite sums $\sum_{n=1}^{N} c_{\alpha_n} e_{\alpha_n}$ that all belong to the algebraic span of $(e_\alpha)_{\alpha\in A}$, which gives the result. Conversely, the algebraic span of $(e_\alpha)_{\alpha\in A}$ is a subset of the Hilbert space span of $(e_\alpha)_{\alpha\in A}$, and the Hilbert space span of $(e_\alpha)_{\alpha\in A}$ is closed, so it must be the entire space H.

To prove that (i) implies (iv), notice that we may write any element $x \in H$ as a countable sum $\sum_{n=1}^\infty c_{\alpha_n} e_{\alpha_n}$. Then we may compute just like in exercise 1.4.18(iv) to obtain the identity

$$x = \sum_{n=1}^\infty \langle x, e_{\alpha_n}\rangle e_{\alpha_n} = \sum_{\alpha\in A} \langle x, e_\alpha\rangle e_\alpha$$

as needed. Similarly, an easy calculation proves that (iv) implies (iii).

We prove that (iii) implies (v). Suppose $x \in H$ is orthogonal to all vectors e_α. Then $\langle x, e_\alpha\rangle = 0$ for $\alpha \in A$, and so $\|x\|^2 = \sum_{\alpha\in A} |\langle x, e_\alpha\rangle|^2 = 0$, which implies that $x = 0$ as needed.

Now we prove that (v) implies (i). Suppose $x \in H$ is orthogonal to the Hilbert space span of $(e_\alpha)_{\alpha\in A}$. Then $(x, e_\alpha) = 0$ for all $\alpha \in A$, and so $x = 0$ by hypothesis. Thus, the orthogonal complement of the Hilbert space span of $(e_\alpha)_{\alpha\in A}$ is trivial, and so the span must be dense in H by exercise 1.4.13. Since it is closed, it must be equal to H.

Now we know that (i)–(v) are equivalent. We prove (i) implies (vi). Let $\phi\colon \ell^2(A) \to H$ be the isometric embedding of $\ell^2(A)$ into H that defines the Hilbert space span, so that $\phi((c_\alpha)_{\alpha\in A}) = \sum_{\alpha\in A} c_\alpha e_\alpha$. Then, given $x \in H$, we may write $x = \sum_{\alpha\in A} c_\alpha e_\alpha$, and so $x = \phi((c_\alpha)_{\alpha\in A})$, which proves that the image of ϕ is H, as needed.

Finally, we prove that (vi) implies (iii). Let $x \in H$. Then $x = \phi((c_\alpha)_{\alpha\in A})$ for some $(c_\alpha)_{\alpha\in A} \in \ell^2(A)$, and we thus compute (using the fact that ϕ is an isometry)

$$\begin{aligned}
\|x\|^2 &= \langle x, x\rangle \\
&= \langle (c_\alpha)_{\alpha\in A}, (c_\alpha)_{\alpha\in A}\rangle \\
&= \sum_{\alpha\in A} |c_\alpha|^2 \\
&= \sum_{\alpha\in A} |\langle (c_\alpha)_{\alpha\in A}, \delta_\alpha\rangle|^2 \\
&= \sum_{\alpha\in A} |\langle x, e_\alpha\rangle|^2.
\end{aligned}$$

Exercise 1.4.20. We use Zorn's lemma to prove that every vector space has an algebraic basis. Consider the poset of linearly independent subsets of a vector space V ordered by inclusion; it is non-empty as it contains the empty set. Then, given a chain, we prove that its union S is linearly independent. Indeed, if $\sum_{i=1}^n c_i v_i$ for $v_i \in S$, then the v_i all belong to some set in the chain, which is linearly independent, and so $c_i = 0$. Thus we obtain a maximal linearly independent set B; we claim that this is an algebraic basis for V. Indeed, suppose for contradiction that $v \in V$ cannot be expressed as a finite linear combination of elements of B. Then it is easy to verify that $B \cup \{v\}$ is linearly independent, which contradicts the maximality of B.

Exercise 1.4.21. Let $\phi\colon \ell^2(A) \to \ell^2(B)$ be an isomorphism. Given an orthonormal basis $(e_\beta)_{\beta\in B}$ for $\ell^2(B)$ (Proposition 1.4.18 guarantees its existence), every basis element e_β may be written as $\phi(\sum_{\alpha\in A} c_\alpha e_\alpha)$

for some element $\sum_{\alpha \in A} c_\alpha e_\alpha \in \ell^2(A)$. This sum must only have at most countably many non-zero terms, so we may write $\sum_{\alpha \in A} c_\alpha e_\alpha = \sum_{n=1}^{\infty} c_{\alpha_n} e_{\alpha_n}$. In this fashion, we obtain a cover of B by a family of at most countable sets indexed by A (namely, $\{e_{\alpha_1}, e_{\alpha_2}, \dots\}$ covers $\{e_\beta\}$). This yields an injection $B \to A$ by the axiom of choice; we may argue similarly to obtain an injection $A \to B$. Therefore, the Schröder–Bernstein theorem implies the existence of a bijection $A \to B$, as needed. (Admittedly, this is overkill for the case where one of the index sets is finite, but it works.)

Exercise 1.4.22. If $(e_\alpha)_{\alpha \in A}$ and $(e_\beta)_{\beta \in B}$ are both orthonormal bases for a Hilbert space H, then we see by exercise 1.4.19 that $\ell^2(A) \approx H \approx \ell^2(B)$. Thus, by exercise 1.4.21, we have $A \approx B$ as needed.

Exercise 1.4.23. Let $(f_n)_{n \in \mathbf{N}}$ be a countable dense subset of H, and let $(e_\alpha)_{\alpha \in A}$ be an orthonormal basis. Then, each f_n may be expressed as a countable sum $\sum_{k=1}^{\infty} c_{n,k} e_{\alpha_{n,k}}$. The collection $(e_{\alpha_{n,k}})_{n,k \in \mathbf{N}}$ is at most countable, although its elements may not be distinct. Write its distinct elements as $(e_n)_{n \in \mathbf{N}}$. We prove that the algebraic span of $(e_n)_{n \in \mathbf{N}}$ is dense in H. Indeed, let $x \in H$, and let $\epsilon > 0$. Then, by density, there exists some f_n with $\|f_n - x\| \leq \epsilon/2$. This f_n is in turn a countable sum $\sum_{k=1}^{\infty} c_{n,k} e_{n,k}$, and so we have

$$\left\| f_n - \sum_{k=1}^{N} c_{n,k} e_{n,k} \right\| \leq \epsilon/2$$

for sufficiently large N. Therefore, the algebraic span of $(e_n)_{n \in \mathbf{N}}$ is dense in H, as needed. In particular, $(e_n)_{n \in \mathbf{N}}$ is an orthonormal basis for H that is a subset of $(e_\alpha)_{\alpha \in A}$, and so it must have been the entirety of $(e_\alpha)_{\alpha \in A}$ to begin with.

Conversely, suppose $(e_\alpha)_{\alpha \in A}$ is an at most countable basis for H. Then, by exercise 1.4.19, the algebraic span of $(e_\alpha)_{\alpha \in A}$ is dense in H. This span might be uncountable, but we may replace the coefficients with rational coefficients (so that they are of the form $p + iq$ with $p, q \in \mathbf{Q}$). This is still dense in H, and is countable, as needed.

Exercise 1.4.24. (Sketch) Let $(e_\alpha)_{\alpha \in A}$ and $(e_\beta)_{\beta \in B}$ be orthonormal bases for H and H'. Then we may construct the tensor product of H and H' as vector spaces as usual. Define on this vector space $H \otimes H'$ an inner product as specified by (ii) and extended by linearity. This space need not be complete with respect to this inner product, so we must take its Hilbert space completion. …

Exercise 1.4.25. [I am quite lost for this and the previous exercise. I have written down some ideas gathered after reading `https://math.stackexchange.com/q/433635/` and `https://math.stackexchange.com/q/2349297/`.] If we are given *countable* orthogonal bases $(\phi_i)_{i \in \mathbf{N}}$ and $(\psi_j)_{j \in \mathbf{N}}$ for $L^2(X)$ and $L^2(Y)$ respectively, then the tensor products $(\phi_i \otimes \psi_j)_{i,j \in \mathbf{N}}$ form an orthogonal basis for $L^2(X \times Y)$. Indeed, we

See also `https://www.ime.usp.br/~tausk/texts/TensorL2.pdf`. It seems that the claim fails if the σ-finite hypothesis isn't present!

may compute

$$\begin{aligned}
\langle \phi_i \otimes \psi_j, \phi_k \otimes \psi_l \rangle_{L^2(X\times Y)} &= \int_{X\times Y} (\phi_i \otimes \psi_j)(x,y)\overline{(\phi_k \otimes \psi_l)(x,y)}\, d(\mu \times \nu) \\
&= \int_{X\times Y} \phi_i(x)\psi_j(y)\overline{\phi_k(x)\psi_l(y)}\, d(\mu\times\nu) \\
&= \int_X \phi_i(x)\overline{\phi_k(x)}\, d\mu \int_Y \psi_j(y)\overline{\psi_l(y)}\, d\nu \\
&= \langle \phi_i, \phi_k\rangle_{L^2(X)}\langle \psi_j, \psi_l\rangle_{L^2(Y)} \\
&= \delta_{ik}\delta_{jl}.
\end{aligned}$$

Now, if $g \in L^2(X \times Y)$ is orthogonal to all $\phi_i \otimes \psi_j$, then

$$\int_X \phi_i(x)\, dx \int_Y \psi_j(y) g(x,y)\, dy = 0$$

for all i, j. Thus the function

$$x \mapsto \int_Y \psi_j(y) g(x,y)\, dy$$

is zero almost everywhere, which in turn implies that g is zero almost everywhere as needed.

One moral of the above story is, of course, that we must be very careful when we give advice to younger people; sometimes they follow it!

— Edsger W. Dijkstra, *The Humble Programmer* (1972)

1.5. Duality and the Hahn–Banach theorem

迷生寂亂 *Rest and unrest derive from illusion;*
悟無好悪 *with enlightenment there is no liking and disliking.*
一切二邊 *All dualities come from*
妄自斟酌 *ignorant inference.*
夢幻虛華 *They are like dreams of flowers in the air:*
何勞把捉 *foolish to try to grasp them.*
得失是非 *Gain and loss, right and wrong:*
一時放卻 *such thoughts must finally be abolished at once.*

—鑑智僧璨,《信心銘》(c. 600)

Exercise 1.5.1. [Had to look up a bit to realize I needed compactness somewhere.] Let $T\colon X \to Y$, and let $\{e_1, \ldots, e_n\}$ be a basis for X. Let $M := \max_{1\le i\le n} \|T(e_i)\|_Y$. Then

$$\|T(v)\|_Y = \Big\|\sum_{i=1}^n v_i T(e_i)\Big\|_Y \le \sum_{i=1}^n |v_i|\|T(e_i)\|_Y \le M\sum_{i=1}^n |v_i|.$$

It remains to show that $\sum_{i=1}^n |v_i| \le C\|v\|_X$. In fact, it suffices to prove the claim for vectors satisfying $\sum_{i=1}^n |v_i| = 1$ by homogeneity, in which case the claim reduces to proving that $C \le \|v\|_X$ for some constant $C > 0$. Consider the set

$$S := \Big\{(c_1, \ldots, c_n) \in \mathbf{C}^n : \sum_{i=1}^n |c_i| = 1\Big\}.$$

(This is the unit sphere in the ℓ^1 norm.) It is closed and bounded, and thus compact. Define $f\colon S \to [0, +\infty)$ by $f(c_1, \ldots, c_n) := \|\sum_{i=1}^n c_i e_i\|_X$. Then f is continuous, as we may compute (with $C := \max_{1\le i\le n} \|e_i\|_X$)

$$\begin{aligned}\Big|\Big\|\sum_{i=1}^n c_i e_i\Big\|_X - \Big\|\sum_{i=1}^n d_i e_i\Big\|_X\Big| &\le \Big\|\sum_{i=1}^n (c_i - d_i)e_i\Big\|_X \\ &\le \sum_{i=1}^n |c_i - d_i|\|e_i\|_X \\ &\le C\sum_{i=1}^n |c_i - d_i| \\ &\le C\sqrt{n}\|(c_1, \ldots, c_n) - (d_1, \ldots, d_n)\|_{\mathbf{C}^n},\end{aligned}$$

where we used the Cauchy–Schwarz inequality at the end, and where $\|\cdot\|_{\mathbf{C}^n}$ denotes the standard Euclidean norm on $\mathbf{C}^n$. By the non-degeneracy of $\|\cdot\|_X$, f is positive everywhere, and so the extreme value theorem gives us a point $s \in S$ for which $f(s) > 0$ and $f(s) = \min_{s'\in S} f(s')$, as needed.

Exercise 1.5.2. We prove that $\|\cdot\|_{\text{op}} := \|\cdot\|_{B(X\to Y)}$ is a norm. If $\|T\|_{\text{op}} = 0$, then $\|Tx\|_Y \le 0\|x\|_X$ for all $x \in X$, so that $Tx = 0$ for all $x \in X$. Thus $T = 0$. Conversely, it is easy to see that $\|0\|_{\text{op}} = 0$.

Next, we consider $\|aT\|_{\text{op}}$. The case for $a = 0$ is clear from non-degeneracy, so suppose that $a \ne 0$. Then

$$\|(aT)x\|_Y = |a|\|Tx\|_Y \le |a|\|T\|_{\text{op}}\|x\|_X$$

for all $x \in X$, so that $\|aT\|_{\mathrm{op}} \leq |a|\|T\|_{\mathrm{op}}$. Similarly, since $\|Tx\|_Y = \frac{1}{|a|}\|(aT)x\|_Y \leq \frac{1}{|a|}\|aT\|_{\mathrm{op}}\|x\|_X$, we have $|a|\|T\|_{\mathrm{op}} \leq \|aT\|_{\mathrm{op}}$ as needed.

Finally, we have

$$\|Sx + Tx\|_Y \leq \|Sx\|_Y + \|Tx\|_Y \leq \|S\|_{\mathrm{op}}\|x\|_X + \|T\|_{\mathrm{op}}\|x\|_X$$

for all $x \in X$, so that $\|S + T\|_{\mathrm{op}} \leq \|S\|_{\mathrm{op}} + \|T\|_{\mathrm{op}}$ as needed.

Now, suppose Y is complete, and let $(T_n)_{n\in\mathbf{N}} \subset B(X \to Y)$ be a Cauchy sequence. Then, given $\epsilon > 0$, there exists N such that $\|T_m - T_n\|_{\mathrm{op}} \leq \epsilon$ whenever $m, n \geq N$. That is, $\|T_m x - T_n x\|_Y \leq \epsilon\|x\|_X$ whenever $m, n \geq N$ and $x \in X$. Thus $(T_n x)_{n\in\mathbf{N}} \subset Y$ is Cauchy for each $x \in X$, and thus converges to a limit $Tx \in Y$. We must prove that $T \in B(X \to Y)$, and that $\lim_{n\to\infty} \|T_n - T\|_{\mathrm{op}} = 0$.

Since Cauchy sequences are bounded, we may choose C with $\|T_n\|_{\mathrm{op}} \leq C$ for $n \in \mathbf{N}$. Then, for large n, we have

$$\begin{aligned}\|Tx\|_Y &\leq \|Tx - T_n x\|_Y + \|T_n x\|_Y \\ &\leq \epsilon + \|T_n\|_{\mathrm{op}}\|x\|_X \\ &\leq \epsilon + C\|x\|_X\end{aligned}$$

for all $x \in X$. Sending $\epsilon \to 0$ proves that $T \in B(X \to Y)$ as needed.

Finally, we prove that $\lim_{n\to\infty} \|T_n - T\|_{\mathrm{op}} = 0$. It suffices to prove that $\|T_n x - Tx\|_Y \leq \epsilon\|x\|_X$ for large n and all $x \in X$. If $\|x\|_X = 0$, this is trivial. If $\|x\|_X = 1$, this is precisely the definition of T — in particular, we defined $Tx := \lim_{n\to\infty} T_n x$, so that $\|T_n x - Tx\|_Y \leq \epsilon$ for large n. Finally, if $\|x\|_X \neq 0$, the result follows from the $\|x\|_X = 1$ case together with homogeneity.

Exercise 1.5.3. We compute

$$\|STx\|_Z \leq \|S\|_{\mathrm{op}}\|Tx\|_Y \leq \|S\|_{\mathrm{op}}\|T\|_{\mathrm{op}}\|x\|_X,$$

which implies that $\|ST\|_{\mathrm{op}} \leq \|S\|_{\mathrm{op}}\|T\|_{\mathrm{op}}$.

Exercise 1.5.4. (Sketch) (i) The construction of the completion is standard (see exercise 1.4.8). The isomorphism is defined in the same way, except now we must prove that the map is an isometry. If $v \in \overline{V}$, then $\phi(v) = \lim_{n\to\infty} v_n$ in $\overline{V}'$, and so

$$\|\phi(v)\|_{\overline{V}'} = \lim_{n\to\infty} \|v_n\|_{\overline{V}'} = \lim_{n\to\infty} \|v_n\|_V = \|v\|_{\overline{V}},$$

where we took the limit in $\overline{V}'$ at the end.

(ii) The map $X^* \to \overline{X}^*$ is defined by extending $f\colon X \to \mathbf{C}$ by density of X in $\overline{X}$ together with continuity; namely, if $x \in \overline{X}$, then we have a sequence $(x_n)_{n\in\mathbf{N}} \subset X$ converging to x, and so we define $f(x) := \lim_{n\to\infty} f(x_n)$. The map $\overline{X}^* \to X^*$ is defined by restriction. Then we may verify by continuity that $\|f\|_{\mathrm{op}} = \|\overline{f}\|_{\mathrm{op}}$.

Exercise 1.5.5. We only prove the case for $\mathbf{C}^n$. Define a map $\mathbf{C}^n \to (\mathbf{C}^n)^*$ by $x \mapsto \langle -, x\rangle$. By the Riesz representation theorem, we obtain an inverse $(\mathbf{C}^n)^* \to \mathbf{C}^n$; thus we have a bijection between $\mathbf{C}^n$ and its dual. It remains to prove that the norm is preserved. Indeed, we have $\|\langle y, x\rangle\|_{\mathbf{C}} \leq \|x\|_{\mathbf{C}^n}\|y\|_{\mathbf{C}^n}$ for all $y \in \mathbf{C}^n$ by the Cauchy–Schwarz inequality; this gives us the bound $\|\langle -, x\rangle\|_{(\mathbf{C}^n)^*} \leq \|x\|_{\mathbf{C}^n}$. Setting $y = x$, we see that this bound is attained, and thus we have

$$\|\langle -, x\rangle\|_{(\mathbf{C}^n)^*} = \|x\|_{\mathbf{C}^n}$$

as needed.

Exercise 1.5.6. We begin with a lemma that we will need for (i). (There is probably a simpler way to do this exercise, but this is what I came up with.)

Lemma. *Let $(c_n)_{n\in\mathbf{N}} \subset \mathbf{C}$ be a sequence of complex numbers, and suppose there exists a constant $C > 0$ such that $|\sum_{n\in S} c_n| \leq C$ for every finite subset $S \subset \mathbf{N}$. Then $(c_n)_{n\in\mathbf{N}}$ is absolutely summable; that is, $\sum_{n\in\mathbf{N}} |c_n| < \infty$.*

Proof. Suppose contrapositively that $\sum_{n\in\mathbf{N}} |c_n| = \infty$. We may split this sum into four sums, depending on which quadrant of the complex plane c_n lies in (we make an arbitrary choice as to which quadrants the axes belong to). Thus we may write the sum as

$$\sum_{\Re(c_n)\geq 0, \Im(c_n)\geq 0} + \sum_{\Re(c_n)\geq 0, \Im(c_n)< 0} + \sum_{\Re(c_n)< 0, \Im(c_n)\geq 0} + \sum_{\Re(c_n)< 0, \Im(c_n)< 0}.$$

One of these sums must be infinite. Suppose it is the first; the other cases are handled similarly. Let $S_{\mathrm{I}} := \{n \in \mathbf{N} : \Re(c_n) \geq 0, \Im(c_n) \geq 0\}$. Then

$$\begin{aligned} \infty &= \sum_{n\in S_{\mathrm{I}}} |c_n| \\ &= \sum_{n\in S_{\mathrm{I}}} \sqrt{\Re(c_n)^2 + \Im(c_n)^2} \\ &\leq \sum_{n\in S_{\mathrm{I}}} \Re(c_n) + \sum_{n\in S_{\mathrm{I}}} \Im(c_n); \end{aligned}$$

thus one of the sums is infinite; say

$$\sum_{n\in S_{\mathrm{I}}} \Re(c_n) = \infty.$$

Then we may choose a large finite subset $S \subset S_{\mathrm{I}}$ for which

$$\sum_{n\in S} \Re(c_n) > C.$$

Thus, we have

$$\Big|\sum_{n\in S} c_n\Big| \geq \Big|\Re \sum_{n\in S} c_n\Big| = \Big|\sum_{n\in S} \Re(c_n)\Big| > C$$

as needed. $\square$

(i) Recall that an isomorphism between normed vector spaces is a continuous invertible linear isometry. Define a linear map

$$\begin{aligned} \phi\colon B(c_c(\mathbf{N}) \to \mathbf{C}) &\longrightarrow \ell^1(\mathbf{N}) \\ f &\longmapsto (f(e_n))_{n\in\mathbf{N}}. \end{aligned}$$

Why is $\phi(f) \in \ell^1(\mathbf{N})$? We have $\|f\|_{\mathrm{op}} < \infty$, and

$$|f((x_n)_{n\in\mathbf{N}})| \leq \|f\|_{\mathrm{op}} \|(x_n)_{n\in\mathbf{N}}\|_{\ell^\infty(\mathbf{N})}$$

for all $(x_n)_{n\in\mathbf{N}} \in c_c(\mathbf{N})$. In particular,

$$\Big|\sum_{n\in S} f(e_n)\Big| = |f((e_n)_{n\in S})| \leq \|f\|_{\mathrm{op}}$$

for all finite subsets $S \subset \mathbf{N}$. Thus, the lemma above implies that $\sum_{n\in\mathbf{N}} |f(e_n)| < \infty$, so that $(f(e_n))_{n\in\mathbf{N}} \in \ell^1(\mathbf{N})$. Define a linear map

$$\begin{aligned} \psi \colon \ell^1(\mathbf{N}) &\longrightarrow B(c_c(\mathbf{N}) \to \mathbf{C}) \\ (a_n)_{n\in\mathbf{N}} &\longmapsto \Big((b_n)_{n\in\mathbf{N}} \mapsto \sum_{n\in\mathbf{N}} a_n b_n\Big). \end{aligned}$$

Then

$$\Big|\sum_{n\in\mathbf{N}} a_n b_n\Big| \leq \sum_{n\in\mathbf{N}} |a_n||b_n| \leq \|(b_n)_{n\in\mathbf{N}}\|_{\ell^\infty(\mathbf{N})} \sum_{n\in\mathbf{N}} |a_n|,$$

and so $\|\psi((a_n)_{n\in\mathbf{N}})\|_{\mathrm{op}} \leq \sum_{n\in\mathbf{N}} |a_n| < \infty$, so that $\psi((a_n)_{n\in\mathbf{N}}) \in B(c_c(\mathbf{N}) \to \mathbf{C})$. To see that this bound is attained, let $(b_n)_{n=1}^\infty := (\overline{a_n}/|a_n|)_{n=1}^N$. Then $\|(b_n)_{n=1}^\infty\|_{\ell^\infty(\mathbf{N})} = 1$, so that

$$\Big|\sum_{n\in\mathbf{N}} a_n b_n\Big| = \Big|\sum_{n=1}^N |a_n|\Big| = \|(b_n)_{n=1}^\infty\|_{\ell^\infty(\mathbf{N})} \sum_{n=1}^N |a_n|,$$

which implies that $\|\psi((a_n)_{n\in\mathbf{N}})\|_{\mathrm{op}} \geq \sum_{n=1}^N |a_n|$. It follows that

$$\|\psi((a_n)_{n\in\mathbf{N}})\|_{\mathrm{op}} = \sum_{n\in\mathbf{N}} |a_n| = \|(a_n)_{n\in\mathbf{N}}\|_{\ell^1(\mathbf{N})},$$

and so ψ is an isometry.

Now, we prove that ϕ and ψ are inverses. We compute

$$\begin{aligned} \phi\psi((a_n)_{n\in\mathbf{N}}) &= \phi\Big((b_n)_{n\in\mathbf{N}} \mapsto \sum_{n\in\mathbf{N}} a_n b_n\Big) \\ &= \bigg(\Big((b_n)_{n\in\mathbf{N}} \mapsto \sum_{n\in\mathbf{N}} a_n b_n\Big)(e_n)\bigg)_{n\in\mathbf{N}} \\ &= (a_n)_{n\in\mathbf{N}} \end{aligned}$$

and

$$\begin{aligned} \psi\phi(f) &= \psi((f(e_n))_{n\in\mathbf{N}}) \\ &= \Big((b_n)_{n\in\mathbf{N}} \mapsto \sum_{n\in\mathbf{N}} f(e_n) b_n\Big) \\ &= \Big((b_n)_{n\in\mathbf{N}} \mapsto f\Big(\sum_{n\in\mathbf{N}} b_n e_n\Big)\Big) \\ &= f. \end{aligned}$$

Thus $\phi = \psi^{-1}$; in particular, ϕ is an isometry, and so ϕ is continuous as needed.

(ii) We first prove that $c_0(\mathbf{N})$ is complete. Suppose $((a_{n,k})_{k\in\mathbf{N}})_{n\in\mathbf{N}}$ is Cauchy, so that

$$\|(a_{m,k})_{k\in\mathbf{N}} - (a_{n,k})_{k\in\mathbf{N}}\|_{\ell^\infty(\mathbf{N})} \leq \epsilon$$

whenever $m, n \geq N$. Then, for all k, we have

$$|a_{m,k} - a_{n,k}| \leq \epsilon \tag{$*$}$$

whenever $m, n \geq N$, so that $(a_{n,k})_{n\in\mathbf{N}} \subset \mathbf{C}$ is Cauchy and converges to a limit $a_k \in \mathbf{C}$.

We prove that $(a_k)_{k\in\mathbf{N}} \in c_0(\mathbf{N})$. Taking the limit $m \to \infty$ in $(*)$, we see that

$$|a_k - a_{n,k}| \leq \epsilon \tag{$\dagger$}$$

for all k, whenever $n \geq N$. Since $(a_{N,k})_{k\in\mathbf{N}} \in c_0(\mathbf{N})$, we have

$$|a_k| \leq |a_k - a_{N,k}| + |a_{N,k}| \leq \epsilon + |a_{N,k}| \leq 2\epsilon$$

whenever k is large. Thus $\lim_{k\to\infty} a_k = 0$ as needed.

Now we prove that $(a_{n,k})_{k\in\mathbf{N}} \to (a_k)_{k\in\mathbf{N}}$ as $n \to \infty$ in $c_0(\mathbf{N})$. We must prove that $\lim_{n\to\infty} \|(a_{n,k})_{k\in\mathbf{N}} - (a_k)_{k\in\mathbf{N}}\|_{\ell^\infty(\mathbf{N})} = 0$. Since

$$\|(a_{n,k})_{k\in\mathbf{N}} - (a_k)_{k\in\mathbf{N}}\|_{\ell^\infty(\mathbf{N})} = \sup_{k\in\mathbf{N}} |a_{n,k} - a_k| \leq \epsilon$$

whenever $n \geq N$ by $(\dagger)$, we are done.

Next, we prove that $c_c(\mathbf{N})$ is dense in $c_0(\mathbf{N})$. Let $(b_n)_{n\in\mathbf{N}} \in c_0(\mathbf{N})$. Then $|b_n| \leq \epsilon$ for $n \geq N$, and $\max\{|b_1|, \ldots, |b_{N-1}|\} < \infty$, so $\|(b_n)\|_{\ell^\infty(\mathbf{N})} < \infty$. In particular, we may consider the truncated sequences $(b_n)_{n=1}^N \in c_c(\mathbf{N})$; thus $\|(b_n)_{n=N+1}^\infty\|_{\ell^\infty(\mathbf{N})} \leq \epsilon$ as needed.

(iii) The proof is similar to (i). Define the linear maps

$$\begin{aligned} \phi\colon B(\ell^1(\mathbf{N}) \to \mathbf{C}) &\longrightarrow \ell^\infty(\mathbf{N}) \\ f &\longmapsto (f(e_n))_{n\in\mathbf{N}} \end{aligned}$$

and

$$\begin{aligned} \psi\colon \ell^\infty(\mathbf{N}) &\longrightarrow B(\ell^1(\mathbf{N}) \to \mathbf{C}) \\ (a_n)_{n\in\mathbf{N}} &\longmapsto \Big((b_n)_{n\in\mathbf{N}} \mapsto \sum_{n\in\mathbf{N}} a_n b_n\Big). \end{aligned}$$

We may verify $\phi \circ \psi = \mathrm{id}_{\ell^\infty(\mathbf{N})}$ and $\psi \circ \phi = \mathrm{id}_{B(\ell^1(\mathbf{N})\to\mathbf{C})}$ as before. We prove that $\phi(f) \in \ell^\infty(\mathbf{N})$. We have

$$|f((c_n)_{n\in\mathbf{N}})| \leq \|f\|_{\mathrm{op}} \|(c_n)_{n\in\mathbf{N}}\|_{\ell^1(\mathbf{N})} = \|f\|_{\mathrm{op}} \sum_{n=1}^\infty |c_n| < \infty$$

for all $(c_n)_{n\in\mathbf{N}} \in \ell^1(\mathbf{N})$. Thus $|f(e_n)| \leq \|f\|_{\mathrm{op}}$ for $n \in \mathbf{N}$, so that $(f(e_n))_{n\in\mathbf{N}} \in \ell^\infty(\mathbf{N})$ as needed.

It remains to prove that ψ is an isometry. We have

$$\begin{aligned} \Big|\sum_{n\in\mathbf{N}} a_n b_n\Big| &\leq \sum_{n\in\mathbf{N}} |a_n||b_n| \\ &\leq \|(a_n)_{n\in\mathbf{N}}\|_{\ell^\infty(\mathbf{N})} \sum_{n\in\mathbf{N}} |b_n| \\ &= \|(a_n)_{n\in\mathbf{N}}\|_{\ell^\infty(\mathbf{N})} \|(b_n)_{n\in\mathbf{N}}\|_{\ell^1(\mathbf{N})}, \end{aligned}$$

so that $\|\psi((a_n)_{n\in\mathbf{N}})\|_{\mathrm{op}} \leq \|(a_n)_{n\in\mathbf{N}}\|_{\ell^\infty(\mathbf{N})}$. To prove the reverse inequality, choose m such that $|a_m| \geq \|(a_n)_{n\in\mathbf{N}}\|_{\ell^\infty(\mathbf{N})} - \epsilon$, and define the sequence $(b_n)_{n=1}^\infty := e_m \in \ell^1(\mathbf{N})$. Then $\|(b_n)_{n\in\mathbf{N}}\|_{\ell^1(\mathbf{N})} = 1$, and

$$\Big|\sum_{n\in\mathbf{N}} a_n b_n\Big| = |a_m| \geq \big(\|(a_n)_{n\in\mathbf{N}}\|_{\ell^\infty(\mathbf{N})} - \epsilon\big) \|(b_n)_{n\in\mathbf{N}}\|_{\ell^1(\mathbf{N})}.$$

Thus ψ is an isometry, and so ϕ is an isometry as well. In particular, ϕ is continuous, and is thus an isomorphism.

Exercise 1.5.7. Let H be a complex vector space, and define the map

$$\begin{aligned} T\colon \overline{H} &\longrightarrow H^* \\ \overline{g} &\longmapsto \langle -, g\rangle_H. \end{aligned}$$

We prove that T is an isomorphism; that is, it is linear, invertible, and an isometry. We have

$$T(\overline{g}+\overline{h}) = T(\overline{g+h}) = \langle -, g+h\rangle_H = \langle -, g\rangle_H + \langle -, h\rangle_H = T(\overline{g}) + T(\overline{h})$$

and

$$T(c\overline{g}) = T(\overline{\overline{c}g}) = \langle -, \overline{c}g\rangle_H = c\langle -, g\rangle_H = cT(\overline{g}),$$

which proves linearity.

To prove invertibility, we use the Hahn–Banach theorem to see that every element of H^* is of the form $\langle -, g\rangle_H$ for some $g \in H$.

Finally, we have $\|\langle -, g\rangle_H\|_{H^*} = \|g\|_H$ by the Cauchy–Schwarz inequality, and

$$\|g\|_H = \langle g, g\rangle_H^{1/2} = \langle \overline{g}, \overline{g}\rangle_{\overline{H}}^{1/2} = \|\overline{g}\|_{\overline{H}},$$

as needed.

Exercise 1.5.8. Consider the map

$$\begin{aligned} T\colon L^{p'}(X, \mathcal{X}, \mu) &\longrightarrow L^p(X, \mathcal{X}, \mu)^* \\ g &\longmapsto \Big(f \mapsto \int_X fg\, d\mu\Big). \end{aligned}$$

By Theorem 1.3.16, there exists a unique $\overline{g} \in L^{p'}$ such that $\lambda = \lambda_{\overline{g}}$, where

$$\lambda_g(f) := \int_X fg\, d\mu;$$

thus T is invertible. Clearly T is linear. By Hölder's inequality, we have $|\int_X fg\, d\mu| \le \|f\|_{L^p}\|g\|_{L^{p'}}$, and so $\|Tg\|_{(L^p)^*} \le \|g\|_{L^{p'}}$. Taking $f := g^{p'-1}$, we see as in the proof of Theorem 1.3.16 that this yields the equality case for Hölder's inequality, so that $\|Tg\|_{(L^p)^*} = \|g\|_{L^{p'}}$ as needed.

Exercise 1.5.9. We compute

$$\|T^*\lambda\|_{X^*} = \|\lambda \circ T\|_{X^*} \le \|\lambda\|_{Y^*}\|T\|_{B(X\to Y)}$$

by exercise 1.5.3. Thus $\|T^*\|_{B(Y^*\to X^*)} \le \|T\|_{B(X\to Y)}$.

Exercise 1.5.10. An $m \times n$ matrix A in $\mathbf{C}^{m\times n}$ may be identified with a linear map $L_A\colon \mathbf{C}^n \to \mathbf{C}^m$, defined by

$$L_A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} A_{11}x_1 + \cdots + A_{1n}x_n \\ \vdots \\ A_{m1}x_1 + \cdots + A_{mn}x_n \end{pmatrix}.$$

Considering the diagram

$$\begin{tikzcd}
\langle -, v\rangle \arrow[rrr, mapsto] & & & \langle L_A(-), v\rangle \\
(\mathbf{C}^m)^* \arrow[r, "*^{-1}"'] \arrow[rrr, bend left, "L_A^*"] & \mathbf{C}^m \arrow[r, "L_{A^t}"'] & \mathbf{C}^n \arrow[r, "*"'] & (\mathbf{C}^n)^* \\
\langle -, v\rangle \arrow[r, mapsto] & v \arrow[r, mapsto] & L_{A^t}(v) \arrow[r, mapsto] & \langle -, L_{A^t}(v)\rangle
\end{tikzcd},$$

we see that it suffices to prove that $\langle L_A(w), v\rangle = \langle w, L_{A^t}(v)\rangle$ for all $v \in \mathbf{C}^m$, $w \in \mathbf{C}^n$. Since

$$\langle L_A(w), v\rangle = \sum_{i=1}^{m} L_A(w)_i v_i = \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij} w_j v_i$$

and

$$\langle w, L_{A^t}(v)\rangle = \sum_{j=1}^{n} w_j L_{A^t}(v)_j = \sum_{j=1}^{n} w_j \sum_{i=1}^{m} A^t_{ji} v_i,$$

the result follows.

Exercise 1.5.11. Suppose $T^*\lambda = 0$, so that $\lambda \circ T = 0$. Given $y \in Y$, we have $y = T(x)$ for some $x \in X$ by surjectivity. Thus $\lambda(y) = \lambda(T(x)) = 0$. It follows that $\lambda = 0$, and so T^* is injective.

If T has a dense image, then given $y \in Y$, we have a sequence $(y_n)_{n\in\mathbf{N}} \subset Y$ converging to y with $y_n = T(x_n)$ by surjectivity. Then continuity of λ and T implies that

$$\lambda(y) = \lim_{n\to\infty} \lambda(y_n) = \lim_{n\to\infty} \lambda(T(x_n)) = \lambda(T(x)) = 0.$$

Exercise 1.5.12. Let $Y \subset X$ be a subspace of a Hilbert space X, and let $\lambda \in Y^*$ with $\|\lambda\|_{\mathrm{op}} = 1$. If Y is not closed, we may extend λ to the closure $\overline{Y}$ by continuity. It is easy to check that the operator norm is preserved. Thus we may assume without loss of generality that Y is closed. By exercise 1.4.7, $\overline{Y}$ is a Hilbert space. We may then apply the Riesz representation theorem to obtain unique $y \in Y$ with $\lambda = \langle -, y\rangle_Y$. We claim that $\widetilde{\lambda} := \langle -, y\rangle_X$ is our desired extension. Clearly this gives a continuous extension of λ. We verify that the operator norm is preserved. By the Cauchy–Schwarz inequality, we have

$$|\langle x, y\rangle|_Y \le \|x\|_Y \|y\|_Y$$

for $x \in Y$; this is an equality iff $x = y$. Since this bound holds for $\langle -, y\rangle_X$ as well, we see that

$$\|\widetilde{\lambda}\|_{X^*} = \|\langle -, y\rangle_X\|_{X^*} = \|y\|_X = \|y\|_Y = \|\langle -, y\rangle_Y\|_{Y^*} = \|\lambda\|_{Y^*} = 1$$

as needed.

Exercise 1.5.13. Note that T being bounded from below implies that it is injective, since then $\|Tx\| = 0$ implies $\|x\| = 0$. Let $\lambda \in X^*$. We must find $\widetilde{\omega} \in Y^*$ for which $T^*\widetilde{\omega} = \lambda$. We begin by defining

$$\begin{aligned} \omega\colon \operatorname{im} T &\longrightarrow \mathbf{C} \\ Tx &\longmapsto \lambda x. \end{aligned}$$

This map is well-defined by injectivity of T. Linearity is easy to verify. Since T is bounded from below, we compute

$$|\lambda x| \leq \|\lambda\|_{X^*}\|x\|_X \leq \frac{\|\lambda\|_{X^*}}{c}\|Tx\|_{Y^*} < \infty.$$

Therefore

$$\|\omega\|_{\mathrm{op}} \leq \frac{\|\lambda\|_{X^*}}{c} < \infty,$$

and so ω is continuous. By the Hahn–Banach theorem, we may extend $\omega\colon \operatorname{im} T \to \mathbf{C}$ to a continuous map $\widetilde{\omega}\colon Y \to \mathbf{C}$, so that $\widetilde{\omega} \in Y^*$. Then

$$T^*\widetilde{\omega} = \widetilde{\omega} \circ T = \omega \circ T = \lambda,$$

as needed.

It is not enough to suppose that T is injective. Consider for example

$$\begin{aligned} T\colon c_0(\mathbf{N}) &\longrightarrow c_0(\mathbf{N}) \\ (a_n)_{n\in\mathbf{N}} &\longmapsto (a_n/n)_{n\in\mathbf{N}}. \end{aligned}$$

It has transpose

$$\begin{aligned} T^*\colon \ell^1(\mathbf{N}) &\longrightarrow \ell^1(\mathbf{N}) \\ (a_n)_{n\in\mathbf{N}} &\longmapsto (T(a_n))_{n\in\mathbf{N}}; \end{aligned}$$

since $\sum_{n\in\mathbf{N}} a_n T(b_n) = T(\sum_{n\in\mathbf{N}} a_n b_n) = \sum_{n\in\mathbf{N}} T(a_n) b_n$, we have:

$$\begin{tikzcd} \Big((b_n)_{n\in\mathbf{N}} \mapsto \sum_{n\in\mathbf{N}} a_n b_n\Big) \arrow[rrr, mapsto] & & & \Big((b_n)_{n\in\mathbf{N}} \mapsto \sum_{n\in\mathbf{N}} a_n T(b_n)\Big) \arrow[dd, mapsto] \\ & c_0(\mathbf{N})^* \arrow[r, "T^*"] & c_0(\mathbf{N})^* \arrow[d, "\approx"] & \\ (a_n)_{n\in\mathbf{N}} \arrow[uu, mapsto] & \ell^1(\mathbf{N}) \arrow[u, "\approx"] & \ell^1(\mathbf{N}) & (T(a_n))_{n\in\mathbf{N}} \end{tikzcd}$$

Since $\sum_{n\in\mathbf{N}} 1/n = \infty$ and $\sum_{n\in\mathbf{N}} 1/n^2 < \infty$, we see that $(1/n^2)_{n\in\mathbf{N}}$ is not in the image of T^*, and so T^* is not surjective.

Exercise 1.5.14. Define $\lambda\colon \operatorname{span}\{x\} \to \mathbf{C}$ by $\lambda(cx) := c\|x\|_X$. Then $\lambda(x) = \|x\|_X$, and $\big|c\|x\|_X\big| = |c|\|x\|_X = \|cx\|$, so that $\|\lambda\|_{\mathrm{op}} = 1$. Linearity is easy to verify. Thus the Hahn–Banach theorem gives us an extension $\widetilde{\lambda}\colon X \to \mathbf{C}$ of λ, with $\|\widetilde{\lambda}\|_{\mathrm{op}} = \|\lambda\|_{\mathrm{op}} = 1$. It follows that $\widetilde{\lambda} \in X^*$ as needed.

Remark. Let $\iota\colon X \to (X^*)^*$ be defined by $x \mapsto \iota(x) := (\lambda \mapsto \lambda(x))$. We show that $\|\iota\|_{\mathrm{op}} \leq 1$. Indeed, when we regard $\iota(x)\colon X^* \to \mathbf{C}$ as the operator, we see that

$$|\lambda(x)| = |\iota(x)(\lambda)| \leq \|\iota(x)\|_{(X^*)^*}\|\lambda\|_{X^*}$$

for all $\lambda \in X^*$. When we treat $\lambda\colon X \to \mathbf{C}$ as the operator, we see that

$$|\lambda(x)| \leq \|\lambda\|_{X^*}\|x\|_X = \|x\|_X\|\lambda\|_{X^*}$$

for all $x \in X$. Therefore, we have

$$\|\iota(x)\|_{(X^*)^*} \leq \|x\|_X$$

for all $x \in X$, and so $\|\iota\|_{\mathrm{op}} \leq 1$ as needed.

Exercise 1.5.15. (i) Suppose $(\lambda_n)_{n\in\mathbf{N}} \subset Y^\perp \subset X^*$ is a sequence of elements converging to $\lambda \in X^*$. Then

$$\iota(y)(\lambda_n) = \lambda_n(y) = 0$$

for all $n \in \mathbf{N}$ and $y \in Y$. Since $\iota(y)$ is continuous, we see that $\lambda(y) = \iota(y)(\lambda) = 0$ for all $y \in Y$, and so $\lambda \in Y^\perp$ as needed. Now we prove that

$$\overline{Y} = \{x \in X : \lambda(x) = 0 \text{ for all } \lambda \in Y^\perp\}.$$

(Here $\overline{Y}$ denotes the closure of Y.) Suppose $y \in \overline{Y}$, so that $y_n \to y$ for some sequence $(y_n)_{n\in\mathbf{N}} \subset Y$. Then, given $\lambda \in Y^\perp$, we have $\lambda(y) = \lim_{n\to\infty} \lambda(y_n) = 0$ as needed. Conversely, ...

(ii) Suppose $Y^\perp$ is trivial. By (i), $\overline{Y} = X$, and so Y is dense. The converse is similar. If Y is trivial, then clearly $Y^\perp = X^*$. If Y is non-trivial, then we may use the Hahn–Banach theorem to produce a functional $\lambda \in X^*$ that is non-zero at some non-zero element $y \in Y$, so that $Y^\perp \neq X^*$.

(iii) ...

(iv) ...

Exercise 1.5.16.

Exercise 1.5.17.

Exercise 1.5.18.

Exercise 1.5.19.

Exercise 1.5.20.

Ask whatever questions you please, but do not ask me for reasons.
A young woman may be forgiven for not being able to give reasons,
since they say she lives in her feelings. Not so with me.
I generally have so many reasons,
and most often such mutually contradictory reasons,
that for this reason it is impossible for me to give reasons.

— Søren Kierkegaard, *Either/Or* **I** (1843)

2.4. *Well-ordered sets, ordinals, and Zorn's lemma*

Exercise 2.4.1. Let $S \subset X$ be the set of elements of X such that $P(x) = \texttt{FALSE}$. We prove that S is empty. Indeed, if S were non-empty, then we may use the well-ordering principle to obtain a minimal element $s \in S$. Then $P(x) = \texttt{TRUE}$ whenever $x < s$ (since any $x < s$ with $P(x) = \texttt{FALSE}$ would violate the minimality of s). By strong induction, it follows that $P(s) = \texttt{TRUE}$, a contradiction.

Let X be a totally ordered set satisfying the principle of strong induction, and let $A \subset X$ be non-empty. Let $L_x := \{a \in A : a \leq x\}$, and define a proposition $P(x)$ that is true iff L_x is either empty or has a least element. Fix $x \in X$, and suppose $P(y)$ is true for all $y < x$; we prove that $P(x)$ holds. First suppose $x \notin A$. Then $L_x = \bigcup_{y<x} L_y$. Either all the L_y in the union are empty, in which case L_x is empty and $P(x)$ holds, or there exists $y < x$ with L_y non-empty. In this case, since $P(y)$ holds by hypothesis, L_y has a least element z. We claim that z is the least element of L_x. Indeed, suppose for contradiction that $z' \in L_x$ satisfies $z' < z$. Then $z' \in A$ with $z' < z \leq a$ for all $a \in L_y$; thus $z' \in L_y$, and so $z \leq z'$, a contradiction.

Now suppose $x \in A$. Then $L_x = \bigcup_{y<x} L_y \cup \{x\}$. If all the sets L_y in the union are empty, then $L_x = \{x\}$, and x is its least element; otherwise the above argument works as before, since x is greater than every element of $\bigcup_{y<x} L_y$.

Therefore, strong induction implies that $P(x)$ holds for $x \in X$. In particular, since $A = \bigcup_{a \in A} L_a$, we may argue as above to see that A has a least element as needed. (Surely there is a simpler argument?)

Exercise 2.4.2. At most one condition can be satisfied, as if $x = \operatorname{succ} y$ for some y, then $\sup([\min(X), x)) = \sup([\min(X), y]) = y \neq x$. Now we argue that at least one condition must be satisfied. Suppose that $x \neq \sup([\min(X), x))$. We prove that $x = \operatorname{succ}(\sup([\min(X), x)))$. Indeed, since

$$\operatorname{succ}\big(\sup([\min(X), x))\big) := \min\big(\big(\sup([\min(X), x)), +\infty\big]\big),$$

it suffices to prove that (i) $x \in \big(\sup([\min(X), x)), +\infty\big]$ and (ii) $x \leq y$ for all $y \in \big(\sup([\min(X), x)), +\infty\big]$. Since $x \neq \sup([\min(X), x))$ and since x is an upper bound for $[\min(X), x)$, (i) follows. If $x > y$, then $y \in [\min(X), x)$. Therefore $y \leq \sup([\min(X), x))$, and so we have $y \notin \big(\sup([\min(X), x)), +\infty\big]$. This proves (ii), and we are done.

Exercise 2.4.3. Suppose $x \neq y$; WLOG $x < y$. We prove that

$$\min((x, +\infty]) < \min((y, +\infty]).$$

Indeed, since $y \in (x, +\infty] \supsetneq (y, +\infty]$, we have

$$\min((x, +\infty]) \leq y < \min((y, +\infty]).$$

Exercise 2.4.4. Let F be the set of all $x \in X$ for which $P(x)$ is false. If F is empty, we are done, so assume it is non-empty and let $m :=$

$\min(F) \in F$ be its minimal element. By exercise 2.4.2, there are two cases to consider.

Limit case. $m = \sup([\min(X), m))$. In this case, $P(y)$ is true for all $y < m$ by minimality of m. Thus the limit case hypothesis for transfinite induction implies that $P(m)$ is true, a contradiction.

Successor case. $m = \operatorname{succ} x$ for some $x \in X$. In this case, $m > x$ and so $P(x)$ is true. The successor case hypothesis in transfinite induction implies that $P(m) = P(\operatorname{succ} x)$ is true, a contradiction.

In either case we obtain a contradiction; thus F must be empty.

Exercise 2.4.5. Let I be an initial segment of X, and consider its complement $X \setminus I$. If $X \setminus I$ is empty, then $I = X = [\min(X), +\infty)$. Otherwise, $X \setminus I$ is non-empty and contains a minimal element m. We claim that $I = [\min(X), m)$. If $x \in I$ with $x \geq m$, then $m \in I$ since I is an initial segment, so $x \in I$ implies $x < m$; thus $I \subset [\min(X), m)$. If $x < m$ then by minimality we have $x \in I$. Thus $[\min(X), m) \subset I$, and we have proven existence.

Uniqueness of m follows from the fact that if $m \neq m'$, say $m < m'$, then $[\min(X), m) \subsetneq [\min(X), m')$, since m is only contained in the latter set.

Exercise 2.4.6. Consider a family $(I_\alpha)_{\alpha \in A}$ of initial segments of X, where we have $I_\alpha = [\min(X), \alpha)$ using exercise 2.4.5 (we discard duplicate sets). We prove that

$$\bigcup_{\alpha \in A} I_\alpha = \bigcup_{\alpha \in A} [\min(X), \alpha) = [\min(X), \sup(A)).$$

The forward inclusion $\subset$ is just the fact that $\alpha \leq \sup(A)$ for $\alpha \in A$. The reverse inclusion $\supset$ follows from the fact that $x < \sup(A)$ implies $x < \alpha$ for some $\alpha \in A$, since $\sup(A)$ is the *least* upper bound of A.

Similarly, we have $\bigcap_{\alpha \in A} I_\alpha = [\min(X), \min(A))$, which can be proven by considering $\bigcap_\alpha I_\alpha = X \setminus (\bigcup_\alpha (X \setminus I_\alpha)) = X \setminus (\bigcup_\alpha [\alpha, +\infty])$ and working as before.

Exercise 2.4.7. By strict monotonicity, we have $\phi([\min(X), \alpha)) \subset [\min(Y), \phi(\alpha))$. Suppose for contradiction that $\phi([\min(X), \alpha)) \neq [\min(Y), \phi(\alpha))$. Then there exists $y < \phi(\alpha)$ such that $y \neq \phi(x)$ for any $x < \alpha$. However, since $\phi(X)$ is an initial segment of Y, we must have $y = \phi(x)$ for some $x \in [\alpha, +\infty)$, contradicting strict monotonicity.

Therefore both $\phi([\min(X), x))$ and $\phi([\min(X), \operatorname{succ} x))$ are initial segments of Y; since

$$[\min(Y), \phi(x)) \cup \{\operatorname{succ} \phi(x)\} = [\min(Y), \phi(\operatorname{succ} x)),$$

it follows that $\phi(\operatorname{succ} x) = \operatorname{succ} \phi(x)$.

Exercise 2.4.8. Let $\phi \colon X \to Y$ be a monotone bijection. Then $\phi(X) = Y$ is an initial segment of Y, and so ϕ is a morphism; we argue similarly for ϕ^{-1}. Conversely, given morphisms $\phi \colon X \to Y$ and $\psi \colon Y \to X$, uniqueness guarantees that $\phi \circ \psi = \mathrm{id}_Y$ and $\psi \circ \phi = \mathrm{id}_X$ (morphisms respect composition by exercise 2.4.7). Since morphisms are monotone, we are done.

Exercise 2.4.9. Let α be the intersection of all ordinals in $\mathcal{F}$ (this might be illegal actually). This is a subset of all ordinals in $\mathcal{F}$. We can also see from the definition that the intersection of ordinals is an ordinal.

Exercise 2.4.10. This should be proven analogously to the corresponding result for well-ordered sets, by considering the least ordinal α for which P fails.

Exercise 2.4.11. Let S be a non-empty set, and fix $s \in S$. Consider the set of well-ordered subsets of S with minimal element s; we may equip this set with a partial order $A \leq B$ if $A \subset B$ and if the well orders are compatible (that is, if $X \subset A \subset B$, then $\min X$ is the same in A and B). Then every chain has an upper bound formed by taking the union, and thus Zorn's lemma implies the existence of a maximal element M. If this were not the full set S we may extend it by appending an element of $S \setminus M$ and declaring it to be larger than all elements of M. Thus $M = S$ and we obtain a well-orderering of S, as needed.

Exercise 2.4.12.

1.6. A quick review of point-set topology

From a five-year-old child to me is only a step.
From the new-born baby to the five-year-old child there is a terrible gap.
From the embryo to the new-born baby there is an abyss.
And from non-existence to the embryo there is not an abyss,
but incomprehensibility.

— LEO N. TOLSTOY, *First Recollections* (1878)

Exercise 1.6.1. (i) Suppose $x_n \to x$, and let U be an open neighborhood containing x. Then $x \in B(x, \epsilon) \subset U$ for some $\epsilon > 0$, and so we may choose N for which $d(x_n, x) < \epsilon$ whenever $n \geq N$; thus $x_n \in U$ whenever $n \geq N$ as needed. Conversely, we obtain the standard definition by considering only the open neighborhoods $B(x, \epsilon)$.

(ii) Suppose x has an open neighborhood U disjoint from E. Then, in particular, there exists $\epsilon > 0$ such that $B(x, \epsilon)$ is disjoint from E. Thus x is not the limit of any sequence of points in E. Conversely, for each n, the open neighborhood $B(x, 1/n)$ intersects E; let x_n be a point of this intersection. It is easy to see that $x_n \to x$.

(iii) Suppose E is closed, and let $x \in X \setminus E$. Then x is not an adherent point of E by definition, and so by (ii) it follows that there exists an open neighborhood $x \in U \subset X \setminus E$; thus $X \setminus E$ is open. Conversely, if $X \setminus E$ is open, then, given an adherent point x of E, we see that every open neighborhood of x intersects E. Thus x cannot belong to $X \setminus E$, and so E contains all of its adherent points as needed.

(iv) Denote by $\bigcap K$ the intersection of all closed sets containing E. Then $\bigcap K$ is closed, and we have $\bigcap K \subset \overline{E}$. Let K be a closed set containing E. We prove that $\overline{E} \subset K$. Indeed, given an adherent point x of E, it is by definition the limit of a sequence in E, which is in turn the limit of a sequence in K. Thus $x \in \overline{K} = K$ as needed.

(v) Suppose E is dense. Then, given a non-empty open set U, it contains some point p, which is necessarily an adherent point of E. Thus $x_n \to p$ for some sequence in E; since U is open, it contains all x_n for sufficiently large n. Thus U intersects E. Conversely, let $x \in X$. Then $B(x, 1/n)$ contains a point $x_n \in E$. Clearly $x_n \to x$; thus x is an adherent point of E.

(vi) Denote by $\bigcup U$ the union of all open sets contained in E. Then $\bigcup U$ is open, and we have $E^\circ \subset \bigcup U$. Let U be an open set contained in E. We prove that $U \subset E^\circ$. Indeed, given $p \in U$, it is an interior point of U, and so it is contained in an open ball B contained in U. This open ball is contained in E, and so p is an interior point of E; that is, $p \in E^\circ$ as needed.

If x is an interior point of E, then there is an open ball containing x that is contained in E. Conversely, if $x \in U \subset E$ for some neighborhood of x, then $x \in B \subset U \subset E$ for some open ball B as needed.

Exercise 1.6.2. This is a standard construction; since I've already worked through it for Hilbert spaces, I will not repeat myself here.

Exercise 1.6.3. Suppose X is complete, and suppose $X \subset Y$. Then, if (x_n) is a convergent sequence in X, it is Cauchy, and thus converges to a limit in X by completeness. Uniqueness of limits thus shows that X is closed as needed.

Conversely, if X is closed in every superspace Y of X, it is in particular closed in its completion $\overline{\mathrm{X}}$. Since it is dense as a subspace of its completion, we see that $X = \overline{\mathrm{X}}$, and so X is complete.

Exercise 1.6.4. Suppose X is totally bounded. Then, there exists a covering of X by n balls of radius 1, with centers $x_1, \ldots, x_n$. Let $D = \max_{i,j} d(x_i, x_j)$; say $d(x_k, x_l) = D$ attains this maximum. Then $X \subset B(x_k, D+1)$, since, given $x \in X$, we have $d(x, x_i) < 1$ for some i, and $d(x_i, x_k) \le D$, so the triangle inequality implies that $d(x_k, x) < D + 1$ as needed.

Suppose $X \subset \mathbf{R}^d$ is bounded. Then $X \subset B(0, M)$ for some M. We prove that $B(0, M)$ may be covered by finitely many balls of radius ϵ. It suffices to cover $[-M, M]^d$ by finitely many balls of radius ϵ. Form a finite lattice

$$\mathcal{L} = \Big\{ (n_1\epsilon/\sqrt{d}, \ldots, n_d\epsilon/\sqrt{d}) \in [-M, M]^d : n_i \in \mathbf{Z} \Big\}.$$

Place balls of radius ϵ at each point of $\mathcal{L}$; this covers X as needed.

Consider $\mathbf{Z}$ with the discrete metric. Then $\mathbf{Z} \subset \overline{B(0,1)}$, but given $\epsilon < 1$, we have $B(n, \epsilon) = \{n\}$, and so there cannot be any finite cover of $\mathbf{Z}$ by balls of radius ϵ.

Exercise 1.6.5. (i) implies (ii). Let $\epsilon > 0$, and use (i) to obtain $\delta > 0$. Then there exists N such that $d_X(x_n, x) \le \delta$ whenever $n \ge N$; thus $d_Y(f(x_n), f(x)) \le \epsilon$ for $n \ge N$ as needed.

(ii) implies (iii). Suppose for contradiction that $x \in f^{-1}(V)$ is such that $B_\epsilon x \not\subset f^{-1}(V)$ for all $\epsilon > 0$. For each n, choose a point $x_n \in B_{1/n}(x) \setminus f^{-1}(V)$. Then $x_n \to x$, and so $f(x_n) \to f(x)$ by (ii). But $f(x_n) \notin V$ for all n, and so x is not an interior point of V. Thus V is not open, a contradiction.

(iii) is equivalent to (iv). The equivalence follows from how $X \setminus f^{-1}(S) = f^{-1}(Y \setminus S)$, and that the complement of an open set is closed (and vice versa).

(iii) implies (i). Let $x \in X$ and $\epsilon > 0$. Then $f^{-1} B_\epsilon f x$ is open by (iii), so $B_\delta x \subset f^{-1} B_\epsilon f x$ for some δ. Thus $d_Y(fx', fx) < \epsilon$ whenever $d_X(x', x) < \delta$, as needed.

Exercise 1.6.6. Suppose $f \oplus g \colon X \to Y \times Z$ is continuous. Then f is continuous, since if $x_n \to x$, then $(fx_n, gx_n) \to (fx, gx)$; thus $fx_n \to fx$. Similarly we see that g is continuous. (Recall that $Y \times Z$ is equipped with the product metric $d_Y \times d_Z((y,z),(y',z')) := \max(d_Y(y,y'), d_Z(z,z'))$.) The converse is similar. Finally, we show that the projection $\pi_Y \colon Y \times Z \to Y$ is continuous. If $(y_n, z_n) \to (y, z)$, then $\max(d_Y(y_n, y), d_Z(z_n, z)) \to 0$. In particular, $d_Y(y_n, y) \to 0$, and we are done.

Exercise 1.6.7. Let $K \subset X$ be compact, and let $(U_\alpha)_{\alpha \in A}$ be an open cover for $f(K)$. Then $(f^{-1}(U_\alpha))_{\alpha\in A}$ is an open cover for K, and thus admits a finite subcover $(f^{-1}(U_{\alpha_i}))_{1\le i\le n}$. Therefore $(U_{\alpha_i})_{1\le i\le n}$ is a finite subcover of $(U_\alpha)_{\alpha\in A}$ as needed.

Exercise 1.6.8. The only hard implication is (iii) $\implies$ (i). Suppose $x_n \not\to x$. Then, there exists open $U \ni x$ such that, for every N, there exists $n \ge N$ with $x_n \ni U$. This yields a subsequence x_{n_i} that lies outside U; none of its subsequences can converge to x as a result.

Exercise 1.6.9. Suppose $x_n \in E$ with $x_n \to x$. Then, given open $U \ni x$, we have $x_n \in U$ for some n by definition of convergence; thus U intersects E as needed.

Suppose $f\colon X \to Y$ is continuous, and $x_n \to x$. Then, given open $V \subset Y$ such that $f(x) \in V$, we see that $f^{-1}(V)$ is open, and so there exists N such that $x_n \in f^{-1}(V)$ whenever $n \ge N$. It follows that $f(x_n) \in V$ whenever $n \ge N$, as needed.

Exercise 1.6.10. The first uncountable ordinal ω_1 can be endowed with the order topology to form a topological space; it is often written $\omega_1 = [0, \omega_1)$. Consider a sequence (α_n) in $[0, \omega_1)$. Then $\alpha = \bigcup_n \alpha_n$ is a countable union of countable ordinals, and thus itself belongs to $[0, \omega_1)$. Since this is the $\limsup$ of (α_n), we may find a subsequence converging to α. Thus $[0, \omega_1)$ is sequentially compact.

The open cover $\{[0, \alpha) : \alpha < \omega_1\}$ has no finite subcover. Thus $[0, \omega_1)$ is not compact.

Exercise 1.6.11. Given $a \ne b$, say $a < b$, we have the disjoint neighborhoods $[a, b)$ and $[b, b+1)$ of a and b respectively; thus $\mathbf{R}$ with the half-open topology is Hausdorff. Suppose for contradiction that the half-open topology is metrizable. Then $(\mathbf{R}, \mathcal{F}_r)$ is separable since $\mathbf{Q}$ is dense. Suppose we had a countable set of basis elements $[a_i, b_i)$. Then, given $x \ne a_i$, the set $[x, x+1)$ is not the union any collection of the sets $[a_i, b_i)$; thus we do not have secound countability. Since second countability is equivalent to separability in metric spaces, this gives the claim.

Exercise 1.6.12. Suppose for contradiction that $(x_n)_{n\in\mathbf{N}}$ is a sequence in a Hausdorff space converging to distinct limits $x \ne y$; choose disjoint neighborhoods V_x and V_y of x and y. Then for sufficiently large n, $x_n \in V_x \cap V_y$, which is absurd.

Exercise 1.6.13. Suppose x is adherent to E, and consider the net $(x_U)_{U\in N(x)}$, where $(N(x), \supset)$ is the directed set of neighborhoods of x ordered by reverse set inclusion, and x_U is an element of $E \cap U$. This net is in E, and it converges to x, since given a neighborhood V of x, we have $x_U \in V$ whenever $U \subset V$. Conversely, let $(x_\alpha)_{\alpha\in A}$ be a net in E that converges to x, and let V be a neighborhood of x. Then $x_\alpha \in V$ for sufficiently large α. Since $x_\alpha \in E$, it follows that $V \cap E$ is non-empty and so x is adherent to E as needed.

Exercise 1.6.14. Let $f\colon X \to Y$ be continuous, and let $(x_\alpha)_{\alpha\in A}$ be a net in X converging to a limit $x \in X$. Then, given a neighborhood V of $f(x)$, continuity gives us an open set $U \ni x$ with $f(U) \subset V$. By convergence of $(x_\alpha)_{\alpha\in A}$, we have $x_\alpha \in U$ for sufficiently large α. Thus $f(x_\alpha) \in V$ for sufficiently large α, so that $(f(x_\alpha))_{\alpha\in A}$ converges to $f(x)$.

Conversely, suppose f is discontinuous, and let $V \subset Y$ be open with $f^{-1}(V) \subset X$ not open. Let x be a non-interior point of $f^{-1}(V)$, and for each neighborhood U of x, choose $x_U \in U \setminus f^{-1}(V)$. Then $(x_U)_{U\in N(x)}$ is a net converging to x, since given a neighborhood W of x, we have $x_U \in W$ whenever $U \subset W$. Thus $(f(x_U))_{U\in N(x)}$ converges to $f(x)$ by hypothesis. In particular, there exists $W \in N(x)$ such that $f(x_U) \in V$ whenever $U \subset W$, and so $f(x_W) \in V$. But this is absurd, since $x_W \notin f^{-1}(V)$ by definition. Therefore f is continuous.

Exercise 1.6.15. [I learned the following from `https://ncatlab.org/nlab/show/compact+spaces+equivalently+have+converging+subnet+of+every+net`. Here I have rewritten it for my own learning.]

Suppose X is compact, and let $(x_\alpha)_{\alpha\in A}$ be a net in X. Given $\alpha \in A$, we define the sets[3] $\alpha^\uparrow := \{x_\beta \in X : \beta \geq \alpha\} \subset X$. Since A is a directed set, $(\alpha^\uparrow)_{\alpha\in A}$ satisfies the finite intersection property (FIP). Thus the collection of closures $(\overline{\alpha^\uparrow})_{\alpha\in A}$ has the FIP as well, and so compactness gives us an element $x \in \bigcap_{\alpha\in A} \overline{\alpha^\uparrow}$.

[3] Strictly speaking, the notation $\alpha^\uparrow$ tends to be reserved for the *upward closure* $\{\beta \in A : \beta \geq \alpha\}$ of α.

We will construct a subnet $(x_{\phi(\beta)})_{\beta\in B}$ of $(x_\alpha)_{\alpha\in A}$ that converges to x. We define the directed set

$$B := \{(\alpha, U) : x_\alpha \in U\} \subset A \times N(x) = (A, \leq) \times (N(x), \supset),$$

where the preorder is given by $(\alpha, U) \leq (\alpha', U')$ iff $\alpha \leq \alpha'$ and $U \supset U'$.

Let us prove that B is a directed set. If we are given (α, U) and (α', U'), we may find an upper bound $\alpha'' \geq \alpha, \alpha'$ in the directed set A. Since $x \in \bigcap_{\alpha\in A} \overline{\alpha^\uparrow}$, the neighborhood $U \cap U'$ of x intersects $\overline{\alpha''^\uparrow}$ and thus[4] intersects $\alpha''^\uparrow$. It follows that there exists $\beta \geq \alpha''$ such that $x_\beta \in U \cap U'$. Therefore, $(\beta, U \cap U')$ is our desired upper bound.

[4] If U is open and C is closed, then $U \cap C \neq \varnothing$ implies $U \cap C^\circ \neq \varnothing$, where C° is the *interior* of C. This follows from the fact that the interior of a finite intersection is equal to the finite intersection of interiors: $(A \cap B)^\circ = A^\circ \cap B^\circ$.

Now we may define $\phi\colon B \to A$ by $\phi((\alpha, U)) := \alpha$. Clearly this is a monotone map, and since it is surjective, it has cofinal image. This completes the construction of the subnet. It remains to prove that $(x_{\phi(\beta)})_{\beta\in B}$ converges to x. Let U be a neighborhood of x. Then U intersects some $\alpha^\uparrow$ as before, and so $x_\beta \in U$ for some $\beta \geq \alpha$. Thus, if $(\beta', V) \geq (\beta, U)$, then $x_{\beta'} \in V \subset U$ as needed.

We now prove the other implication. Suppose that X is not compact. We will construct a net in X with no convergent subnet. By the FIP formulation of compactness, there exists a collection of closed subsets $(C_i)_{i\in I}$ of X that has non-empty finite intersections but has empty intersection $\bigcap_{i\in I} C_i = \varnothing$. Let $\mathrm{Fin}(I)$ be the set of finite subsets of I. Then, for each $J \in \mathrm{Fin}(I)$, we may choose some $x_J \in \bigcap_{i\in J\subset I} C_i$. This defines a net $(x_J)_{J\in\mathrm{Fin}(I)}$, since $(\mathrm{Fin}(I), \subset)$ is a directed set. Suppose for contradiction that this net has a convergent subnet $(x_{\phi(\beta)})_{\beta\in B}$ with $\phi\colon B \to \mathrm{Fin}(I)$ monotone and cofinal; say this subnet converges to $x \in X$. Since $\bigcap_{i\in I} C_i = \varnothing$, there exists $i_x \in I$ such that $x \notin C_{i_x}$. Since C_{i_x} is closed, there exists an open neighborhood U of x that is disjoint from C_{i_x}. Thus $x_J \notin U$ for all $\{i_x\} \subset J \in \mathrm{Fin}(I)$. Now, since ϕ is cofinal, there exists $\beta \in B$ such that $\{i_x\} \subset \phi(\beta)$. Since $(x_{\phi(\beta)})_{\beta \in B}$ converges, there exists β' such that $x_{\phi(\alpha)} \in U$ whenever $\alpha \geq \beta'$. Taking an upper bound $\beta'' \geq \beta, \beta'$, we see that $x_{\phi(\beta'')} \in U$ and $\{i_x\} \subset \phi(\beta') \subset \phi(\beta'')$. But the latter implies that $x_{\phi(\beta'')} \notin U$, which gives the desired contradiction and completes the proof.

Exercise 1.6.16. Let X be Hausdorff and suppose for contradiction that $(x_\alpha)_{\alpha \in A}$ is a net that converges to distinct limits $x \neq x'$. Then, there exists $\beta \in A$ such that for every neighborhood V of x, we have $x_\alpha \in V$ whenever $\alpha \geq \beta$, and similarly there exists $\beta' \in A$ such that for every neighborhood V' of x', we have $x_\alpha \in V'$ whenever $\alpha \geq \beta'$. Let β'' be an upper bound for β and β', and let $V \ni x$ and $V' \ni x'$ be disjoint open sets. Then $x_\alpha \in V \cap V' = \varnothing$ whenever $\alpha \geq \beta''$, which is absurd.

Conversely, suppose X is not Hausdorff, and let $x \neq y$ be distinct inseparable points. Then we may define a net $(x_{(U,V)})_{(U,V) \in N(x) \times N(y)}$ by choosing a point $x_{(U,V)} \in U \cap V$; one may verify that this net converges to both x and y.

Exercise 1.6.17.

1.7. The Baire category theorem and its Banach space consequences

Je continue à être peu brillant,
à suivre une courbe infiniment sinueuse,
sans dérivée peut-être,
mais peu importe.

— René Baire, letter to Émile Borel (1902)

Exercise 1.7.1. Given a sequence $(U_n)_{n\in\mathbf{N}}$ of open dense sets, we have $\bigcap_n U_n = X \setminus \bigcup_n (X \setminus U_n)$ by De Morgan's identity. Each set $X \setminus U_n$ is closed and nowhere dense. Indeed, if $X \setminus U_n$ had nonempty interior, then it contains an open ball B; thus U_n does not contain B, contradicting density. Thus the Baire category theorem implies that $\bigcup_n (X \setminus B_n)$ contains no balls as needed. Now, if a set S is not dense, then $\overline{S} \neq X$, so $X \setminus \overline{S}$ is open, and thus contains a ball. Since $X \setminus \overline{S} \subset X \setminus S$, it follows that $X \setminus S$ contains a ball. In particular, we may conclude that $\bigcap_n U_n$ is dense as needed. The converse is proven similarly.

Exercise 1.7.2. Recall that one of the formulations of density in a metric space is as follows: S is said to be *dense in B* if for every $y \in B$ and $\epsilon > 0$, the intersection $B(y,\epsilon) \cap S$ is non-empty.

Given $x \in X$, the set $\{x\}$ is nowhere dense. Indeed, if $\{x\}$ were dense in a ball B, then, given $y \in B$, if $y \neq x$, then setting $\epsilon := d(x,y)$ shows that the intersection $B(y,\epsilon) \cap \{x\}$ is empty; thus we must have $B = \{x\}$, which implies that x is isolated.

Thus $X = \bigcup_{x\in X}\{x\}$ cannot contain a ball by the Baire category theorem, which is absurd.

Exercise 1.7.3. [I could not solve this problem. I have nothing to add to the fantastic solution given at `https://math.stackexchange.com/a/466343/`.]

Exercise 1.7.4.

Exercise 1.7.5.

Exercise 1.7.6.

Exercise 1.7.7.

Exercise 1.7.8.

Exercise 1.7.9.

Exercise 1.7.10.

1.8. Compactness in topological spaces

Now compactness is a topological property, so to use it,
you really should say explicitly what the topological space is,
and what the open and closed sets are. But mathematicians rarely,
if ever, do that. In fact, they usually don't specify anything at all about
the setting; they just say "by the usual compactness argument" and
move on. That's great for experts, but not so great for beginners.

— Jeffrey Shallit (2013)

Exercise 1.8.1. Given an open cover $(U_\alpha)_{\alpha \in A}$ of a finite set $X = \{x_1, \ldots, x_n\}$, for each $1 \leq i \leq n$ we may choose U_{α_i} containing x_i, so that $(U_{\alpha_i})_{1 \leq i \leq n}$ is a finite subcover.

Let $\bigcup_{1 \leq i \leq n} K_i$ be a finite union of compact subsets of a topological space X, and let $(U_\alpha)_{\alpha \in A}$ be an open cover of $\bigcup_{1 \leq i \leq n} K_i$. Then, for each $1 \leq i \leq n$, $(U_\alpha)_{\alpha \in A}$ covers K_i, and so compactness gives us a finite subcover $(U_{\alpha_{i,j}})_{1 \leq j \leq n_i}$. Taking these covers together, we obtain a finite subcover $(U_{\alpha_{i,j}})_{1 \leq i \leq n; 1 \leq j \leq n_i}$ for $\bigcup_{1 \leq i \leq n} K_i$.

Let $f \colon K \to Y$ be a continuous map from a compact space K to an arbitrary topological space Y. Suppose we are given an open cover $(U_\alpha)_{\alpha \in A}$ for $f(K)$. Then, by continuity, $(f^{-1}(U_\alpha))_{\alpha \in A}$ is an open cover for K; thus we obtain a finite subcover $(f^{-1}(U_{\alpha_i}))_{1 \leq i \leq n}$ by compactness. It follows that $(U_{\alpha_i})_{1 \leq i \leq n}$ is our desired finite subcover for $f(K)$.

Now we prove the corresponding results for sequential compactness. Suppose X is a finite set, and consider a sequence $(x_n)_{n \in \mathbf{N}}$. The pigeonhole principle tells us that there is a constant subsequence $(x_{n_i})_{i \in \mathbf{N}}$; it is thus convergent and we are done.

Let $(x_n)_{n \in \mathbf{N}}$ be a sequence in $\bigcup_{1 \leq i \leq n} K_i$. Then the pigeonhole principle gives us a subsequence that is completely contained in one of the K_i; we may use the sequential compactness of K_i to pass to a convergent subsequence as needed.

Let $f \colon K \to Y$ be a continuous map from a sequentially compact space K to an arbitrary topological space Y, and consider a sequence $(y_n)_{n \in \mathbf{N}}$ in $f(K)$. By the axiom of countable choice, we may choose a sequence $(x_n)_{n \in \mathbf{N}}$ such that $f(x_n) = y_n$. We may then pass to a convergent subsequence $(x_{n_i})_{i \in \mathbf{N}}$; continuity implies that $(y_{n_i})_{i \in \mathbf{N}}$ is convergent as needed.

Exercise 1.8.2. Let $C \subset X$ be a closed subset of a compact space X, and let $(U_\alpha)_{\alpha \in A}$ be an open cover for C. Then $(U_\alpha)_{\alpha \in A} \cup \{X \setminus C\}$ is an open cover for X; thus we have a finite subcover $(U_{\alpha_i})_{1 \leq i \leq n}$, perhaps including $X \setminus C$, of X. It follows that $(U_{\alpha_i})_{1 \leq i \leq n}$ is our desired finite subcover.

Any open cover in $(X, \mathcal{F}')$ is also an open cover in $(X, \mathcal{F})$; thus we may pass to a finite subcover by compactness in $(X, \mathcal{F})$.

The empty set always has the empty subcover. Given a non-empty set $A \subset X$, where X is equipped with the trivial topology, any open cover of A must contain the open set X; then $\{X\}$ is our desired finite subcover for A.

Exercise 1.8.3. Let $K \subset X$ be a compact subset of a Hausdorff space X, and suppose $x \notin K$. Then, for each $y \in K$ there exist disjoint open sets $V_{x,y} \ni x$ and $V_{y,x} \ni y$; thus $(V_{y,x})_{y \in K}$ is an open cover for K and we may obtain a finite subcover $(V_{y_i,x})_{1 \le i \le n}$; it follows that $\bigcup_{1 \le i \le n} V_{y_i,x} \supset K$ and $\bigcap_{1 \le i \le n} V_{x,y_i} \ni x$ are disjoint open sets; in particular, $\bigcap_{1 \le i \le n} V_{x,y_i}$ is an open neighborhood of x disjoint from K. Thus x is not an adherent point of K; it follows that K is closed as needed.

Stronger topologies have more open sets; in particular they contain all the open sets guaranteed by the Hausdorff axiom in the original topology and so they must also be Hausdorff.

In the discrete topology, given $x \neq y$, the sets $\{x\}$ and $\{y\}$ are disjoint open sets as needed.

Exercise 1.8.4. Fix a filter p and consider the set of all filters containing p ordered by inclusion; this set is non-empty as it contains p, and every chain has an upper bound given by its union, which is easily verified to also be a filter. Thus, Zorn's lemma gives us a maximal element p'. Notice that p' is non-empty, since either p is non-empty, or p is empty, in which case $\{X\}$ is a filter on X. Thus we must have $X \in p'$. Suppose $E \subset X$ is such that $E \notin p'$ and $X \setminus E \notin p'$, so that E is non-empty. Then we may define

$$p'' := p' \cup \{A \cap E : A \in p'\} \cup \{B : A' \cap E \subset B \subset X \text{ for some } A' \in p'\}.$$

We claim that this is a filter strictly larger than p'. We have closure under finite intersection by considering six cases (i, j) with $1 \le i \le j \le 3$, depending on which of the three sets in the union to which the sets we are considering belong; for example, in case $(1, 3)$, if $A \in p'$ and $B \supset A' \cap E$, then $(A \cap A') \cap E \subset A \cap B$, and so $A \cap B \in p''$ by monotonicity. In case $(2, 3)$, if $A, A' \in p'$ so that $A \cap E \in p''$ and $B \supset A' \cap E$, then $(A \cap A') \cap E \subset (A' \cap E) \cap B$, and so $(A' \cap E) \cap B \in p''$ by monotonicity. In case $(3, 3)$, if $A', A'' \in p'$ so that $A' \cap E \subset B'$ and $A'' \cap E \subset B''$, then $(A' \cap A'') \cap E \subset B' \subset B''$, and so $B' \cap B'' \in p''$ by monotonicity. Monotonicity of p'' is straightforward from our definition of p''. Finally, if $\varnothing \in p''$, then $A \cap E = \varnothing$ for some $A \in p'$, so that $A \subset X \setminus E$. Monotonicity then implies $X \setminus E \in p'$, contradicting our hypotheses. Thus, we obtain a contradiction to the maximality of p'.

Similarly, if $E \in p'$ and $X \setminus E \in p'$, then their intersection $\varnothing$ also lies in p', a contradiction. It follows that every set $E \subset X$ is such that exactly one of E and $X \setminus E$ lies in p', and so p' is an ultrafilter, as needed.

Exercise 1.8.5. By definition, any filter satisfies the finite intersection property. Conversely, given a collection $\mathcal{C} \subset 2^X$ of subsets of X satisfying the finite intersection property, then clearly $\mathcal{C}$ does not contain the empty set. We define $p := \{A : C \subset A \subset X \text{ for some } C \in \mathcal{C}\}$; it is easy to see that p is a filter containing $\mathcal{C}$.

Exercise 1.8.6. Suppose X is Hausdorff, and suppose p is an ultrafilter converging to distinct points x and y. Then we have disjoint open sets $V_x \ni x$ and $V_y \ni y$ with $V_x, V_y \in p$; it follows that $\varnothing = V_x \cap V_y \in p$, a contradiction.

Conversely, suppose X is not Hausdorff, so that there exist points x and y that cannot be separated by open sets. Thus every neighborhood of x is a neighborhood of y, and vice versa — every finite intersection of such neighborhoods contains x and y, and so the set of neighborhoods of x and y satisfies the finite intersection property. By exercise 1.8.5, this set is contained in an ultrafilter p; thus p converges to both x and y as needed.

Exercise 1.8.7. Suppose X is compact, and consider an ultrafilter p on X. Let $\overline{p} := \{\overline{A} : A \in p\}$. Then $\overline{p}$ is a collection of closed sets satisfying the finite intersection property, and thus compactness guarantees that the intersection $\bigcap \overline{p}$ is nonempty. Let $x \in \bigcap \overline{p}$, and consider a neighborhood U of x. If $\overline{X \setminus U} = X \setminus U \in p$, then $x \in \bigcap \overline{p} \subset X \setminus U$, which is absurd. Thus $U \in p$ as needed, and we conclude that p converges to x as needed.

Conversely, suppose X is not compact, and let $\mathcal{C}$ be a collection of closed subsets of X satisfying the finite intersection property such that $\bigcap \mathcal{C} = \varnothing$. By exercise 1.8.5, $\mathcal{C}$ is contained in an ultrafilter p. Suppose for contradiction that x is a limit of p, so that every neighborhood U of x is contained in p. In particular, given $C \in \mathcal{C}$, if $x \notin C$, the set $X \setminus C$ would be an open neighborhood of x, and so $X \setminus C$ and C would both belong to p, contradicting the fact that p is an ultrafilter. Thus $x \in C$ for all $C \in \mathcal{C}$, contradicting the fact that $\bigcap \mathcal{C} = \varnothing$. This completes the proof.

Exercise 1.8.8. Suppose $\mathcal{B}$ is a base for a topology $\mathcal{F}$. Then $\mathcal{B}$ covers X, since X is open and thus may be expressed as a union of open sets in $\mathcal{B}$. Let $x \in X$, and let $U, V \in \mathcal{B}$ be basic open neighborhoods of x. Then $U \cap V$ is open and thus may be expressed as a union $\bigcup_\alpha B_\alpha$ of basic open sets. Since $U \cap V$ contains x, it follows that $x \in B_\alpha \subset U \cap V$ for some α as needed.

Conversely, suppose $\mathcal{B}$ is a collection of subsets of X that covers X and satisfies the property given in the exercise. Then we may define the topology $\mathcal{F}$ as the collection of all unions of basic open sets of $\mathcal{B}$. The empty union gives the empty set, so $\varnothing \in \mathcal{F}$. The basic open sets cover X by hypothesis, so $X \in \mathcal{F}$. The union of a union remains a union, so $\mathcal{F}$ is closed under arbitrary unions. Finally, since

$$\Big(\bigcup_{\alpha \in A} B_\alpha\Big) \cap \Big(\bigcup_{\alpha' \in A'} B_{\alpha'}\Big) = \bigcup_{(\alpha,\alpha') \in A \times A'} B_\alpha \cap B_{\alpha'},$$

and since the intersection of two basic open sets is itself a union of basic open sets of the property given in the exercise, it follows that $\mathcal{F}$ is closed under finite intersection and is thus a topology as needed.

Exercise 1.8.9. Suppose every basic open cover has a finite subcover, and let $(U_\alpha)_{\alpha \in A}$ be an open cover for X. Write $U_\alpha = \bigcup_{\beta \in A_\alpha} U_{\alpha,\beta}$ with

$U_{\alpha,\beta} \in \mathcal{B}$. Then $(U_{\alpha,\beta})_{\alpha \in A; \beta \in A_\alpha}$ is a basic open cover for X, and thus admits a finite subcover $(U_{\alpha_i,\beta_i})_{1 \le i \le n}$. Since $U_{\alpha_i,\beta_i} \subset U_{\alpha_i}$, it follows that $(U_{\alpha_i})_{1 \le i \le n}$ is our desired finite subcover.

Exercise 1.8.10. Suppose $\mathcal{B}$ is a subbase for $(X, \mathcal{F})$. Then every element of $\mathcal{B}$ is open in X by definition, and so every element of $\mathcal{B}^*$ is open in X as well. Since $\mathcal{B}^*$ is closed under finite intersections, it follows from exercise 1.8.8 that $\mathcal{B}^*$ is a base for $(X, \mathcal{F})$.

Conversely, if $\mathcal{B}^*$ is a base for $(X, \mathcal{F})$, then $\mathcal{B} \subset \mathcal{F}$, and so it suffices to prove that $\mathcal{F}$ is the weakest topology with this property. Let $U \in \mathcal{F}$. Then $U = \bigcup_{\alpha \in A} B_\alpha$, where $B_\alpha \in \mathcal{B}^*$. Thus, if the elements of $\mathcal{B}$ are open, then every set $U \in \mathcal{F}$ must be open as well. This completes the proof.

Exercise 1.8.11. (i) Suppose that $x_n \to x$. Then, every open neighborhood of x contains x_n for sufficiently large n; in particular, this holds for subbasic open neighborhoods of x. Conversely, let U be a neighborhood of x. Then it is the union of finite intersections of subbasic open sets. Consider the finite intersection $B_1 \cap \cdots \cap B_k$ in this union that contains x. Then for sufficiently large n, x_n will be contained in each B_i, and so will be contained in the finite intersection $B_1 \cap \cdots \cap B_k \subset U$ as needed.

(ii) The forward implication is just exercise 1.6.1(ii). For the reverse implication, suppose every basic open neighborhood of x intersects E. Then every neighborhood of x intersects E, and the result again follows from exercise 1.6.1(ii). The result fails for subbases, however. Consider $\mathbf{R}$ with the usual topology, which is generated by the subbase consisting of sets of the form $(-\infty, a)$ and $(b, +\infty)$ for $a, b \in \mathbf{R}$. Every subbasic open neighborhood of 0 then intersects $\mathbf{N} \setminus \{0\}$, but 0 is not an adherent point of this set.

(iii) Suppose x is in the interior of U. Then $x \in V \subset U$ for some open set V. The set V is the union of basic open sets, and so one of these basic open sets contains x as needed. Conversely, suppose U contains a basic open neighborhood B of x. Then B is itself open, and so x is contained in the interior of U. The result fails for subbases — consider $\mathbf{R}$ with subbase as in (ii), with $0 \in (-1, 1)$.

(iv) Suppose the inverse image of every subbasic open set is open, and let $U \subset X$ be an open set. Then $U = \bigcup_{\alpha \in A} B_\alpha$, and $f^{-1}(U) = f^{-1}(\bigcup_{\alpha \in A} B_\alpha) = \bigcup_{\alpha \in A} f^{-1}(B_\alpha)$. The result then follows from the fact that the union of open sets is open.

Exercise 1.8.12. Suppose X is not compact. By exercise 1.8.7, there exists an ultrafilter p with no limits. Then, given $x \in X$, there exists some neighborhood U of x with $U \notin p$. It follows that $x \in U_x \subset U$ for some basic open set U_x, and monotonicity implies that $U_x \notin p$. The set U_x is a finite intersection $B_{x,1} \cap \cdots \cap B_{x,k}$ of subbasic open sets. Since p is closed under finite intersections, one of these subbasic open sets is not in p, say $B_x = B_{x,i}$. Thus we have an subbasic open cover $(B_x)_{x \in X}$ for X. Suppose for contradiction that this cover admits a finite subcover $B_{x_1} \cup \cdots \cup B_{x_k}$. Since each B_{x_i} is not in p, it follows that their union is not in p as well; that is, $X \notin p$, contradicting the fact that p is an ultrafilter. This completes the proof.

Exercise 1.8.13. Consider the subbase for the usual topology on $[0,1]$ consisting of sets of the form $[0,a)$ or $(b,1]$, where $0 \le a,b \le 1$. By the Alexander subbase theorem, it suffices to prove that every subbasic open cover of $[0,1]$ admits a finite subcover. Any such cover is of the form

$$[0,1] = \bigcup_{a\in A}[0,a) \cup \bigcup_{b\in B}(b,1] = [0,\sup A) \cup (\inf B, 1],$$

where $A, B \subset [0,1]$ are non-empty. Thus $\sup A > \inf B$, and so $\sup A \ge a > b \ge \inf B$ for some $a \in A$ and $b \in B$. It follows that $[0,1] = [0,a) \cup (b,1]$ is our desired finite subcover.

Exercise 1.8.14. We first prove that X is Hausdorff. Let $y < z$ in X. If $y < w < z$ for some $w \in X$, then the sets $\{x : x < w\}$ and $\{x : x > w\}$ are open and separate y and z. Otherwise, the sets $\{x : x < z\}$ and $\{x : x > y\}$ are disjoint and thus separate y and z.

Suppose X has no maximal element. Then $(\{x : x < a\})_{a\in X}$ is a subbasic open cover of X, since every element $x \in X$ is smaller than some $a \in X$ by hypothesis. If this cover has a finite subcover, say $\{x : x < a_1\} \cup \cdots \cup \{x : x < a_k\}$, then $\max_{1\le i\le k} a_i$ would be a maximal element for X, a contradiction. Therefore, by the Alexander subbase theorem, X is not compact.

Conversely, suppose X has a maximal element m. Then, given any subset $A \subset X$, we may take the supremum $\sup A$ (since X is well-ordered).[5] We have $\sup A \le m$, and so $\sup A \in X$. (On the other hand, we always have $\inf A \in X$, since we always have the minimal element $\min(X)$ in well-ordered sets X.) Just like in exercise 1.8.13, any subbasic open cover of X is of the form

$$X = \bigcup_{a\in A}\{x \in X : x < a\} \cup \bigcup_{b\in B}\{x \in X : x > b\}$$

for some subsets $A, B \subset X$. Crucially, B is non-empty since if it were, we would have $m \notin X$. Similarly, A is non-empty since if it were, we would have $\min(X) \notin X$. We can rewrite the cover as

$$X = \{x \in X : x < \sup A\} \cup \{x \in X : x > \min B\};$$

thus $\sup A > \min B$ and $\sup A \ge a > b \ge \min B$ for some $a \in A$ and $b \in B$. It follows that $\{x \in X : x < a\} \cup \{x \in B : x > b\}$ is our desired finite subcover, and thus X is compact by the Alexander subbase theorem.

[5] Recall that in well-ordered sets,
$\sup A := \min\{x \in X : x \ge a \text{ for all } a \in A\}.$

Exercise 1.8.15. Let X and Y be Hausdorff spaces, and let $(x,y),(x',y') \in X \times Y$ be distinct points; say $x \ne x'$. Then there exist disjoint open sets $V_x, V_{x'} \subset X$ with $x \in V_x$ and $x' \in V_{x'}$. It follows that $V_x \times Y$ and $V_{x'} \times Y$ are disjoint open neighborhoods of (x,y) and (x',y'), which gives the claim. (The argument works the same way when $y \ne y'$.)

Exercise 1.8.16. Let X and Y be sequentially compact spaces, and let $(x_n, y_n)_{n\in\mathbf{N}}$ be a sequence in $X \times Y$. We may choose a convergent subsequence $x_{n_k} \to x$ of $(x_n)_{n\in\mathbf{N}}$. Then, we may choose a convergent subsequence $y_{n_{k_l}} \to y$ of $(y_{n_k})_{k\in\mathbf{N}}$. Then $(x_{n_{k_l}}, y_{n_{k_l}})_{l\in\mathbf{N}}$ converges to (x, y) as needed.

Exercise 1.8.17.

Exercise 1.8.18. Let $X = \prod_{\alpha\in A} X_\alpha$ be a product of Hausdorff spaces equipped with the product topology, and let $(x_\alpha)_{\alpha\in A}$ and $(x'_\alpha)_{\alpha\in A}$ be distinct points in X; say $x_\beta \neq x'_\beta$ for some $\beta \in A$. Then, since X_β is Hausdorff, there exist disjoint open sets $x_\beta \in V \subset X_\beta$ and $x'_\beta \in V' \subset X_\beta$. It follows that $\pi_\beta^{-1}(V)$ and $\pi_\beta^{-1}(V')$ are open sets in X that separate $(x_\alpha)_{\alpha\in A}$ and $(x'_\alpha)_{\alpha\in A}$, and so X is Hausdorff as needed.

The above proof also shows that X equipped with the box topology is Hausdorff, since the box topology is stronger than the product topology.

Exercise 1.8.19. To verify the triangle inequality for d, it suffices to verify that

$$\frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)} \leq \frac{d_n(x_n, z_n)}{1 + d_n(x_n, z_n)} + \frac{d_n(z_n, y_n)}{1 + d_n(z_n, y_n)}$$

for each n. This is clear if $d_n(x_n, y_n) \geq \max\{d_n(x_n, z_n), d_n(z_n, y_n)\}$; otherwise it follows from the fact that $\alpha \mapsto \alpha/(1+\alpha)$ is increasing for $\alpha \geq 0$.

We must prove that a set $U \subset X$ is open with respect to the metric d if and only if U is open in the product topology. We prove the reverse implication first. Suppose $U \subset X$ is a basic open set in the product topology, and let $(x_n)_{n\in\mathbf{N}} \in U$. We must find $\epsilon > 0$ such that $(y_n)_{n\in\mathbf{N}} \in U$ whenever

$$d((x_n)_{n\in\mathbf{N}}, (y_n)_{n\in\mathbf{N}}) := \sum_{n=1}^{\infty} 2^{-n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)} < \epsilon.$$

Notice that this implies $d_n(x_n, y_n)/(1 + d_n(x_n, y_n)) < 2^n\epsilon$, or

$$d_n(x_n, y_n) < \frac{2^n\epsilon}{1 - 2^n\epsilon}$$

whenever $\epsilon < 2^{-n}$. Since U is basic open, it is of the form $U = \pi_{i_1}^{-1}(U_{i_1}) \cap \cdots \cap \pi_{i_k}^{-1}(U_{i_k})$ for integers $1 \leq i_1 < \cdots < i_k$. Choosing $0 < \epsilon < 2^{-i_k}$ such that

$$B_{d_{i_j}}\left(x_{i_j}, \frac{2^{i_j}\epsilon}{1 - 2^{i_j}\epsilon}\right) \subset U_{i_j}$$

for $1 \leq j \leq k$, we obtain the claim.

Conversely, given $(x_n)_{n\in\mathbf{N}} \in X$ and $\epsilon > 0$, it suffices to prove that there exists a basic open neighborhood of x in the product topology completely contained in the open ball $B_d((x_n)_{n\in\mathbf{N}}, \epsilon)$. Choose large N so that $2^{-N} < \epsilon/2$, and consider the basic open set

$$B := \bigcap_{n=1}^{N} \pi_n^{-1}(B_{d_n}(x_n, \epsilon/2)).$$

Then, for $(y_n)_{n\in\mathbf{N}} \in B$, we have $d_n(x_n, y_n) < \epsilon/2$ for $1 \le n \le N$, and so (since $\alpha \mapsto \alpha/(1+\alpha)$ is increasing for $\alpha \ge 0$)

$$\frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)} < \frac{\epsilon/2}{1+\epsilon/2} < \frac{\epsilon}{2}.$$

We thus compute

$$\begin{aligned} d((x_n)_{n\in\mathbf{N}}, (y_n)_{n\in\mathbf{N}}) &= \Big(\sum_{n=1}^{N} + \sum_{n=N+1}^{\infty}\Big) 2^{-n} \frac{d_n(x_n, y_n)}{1+d_n(x_n, y_n)} \\ &< \sum_{n=1}^{N} 2^{-n} \cdot \frac{\epsilon}{2} + \sum_{n=N+1}^{\infty} 2^{-n} \\ &\le \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon, \end{aligned}$$

which gives the result.

Exercise 1.8.20. Let U be an open neighborhood of $\pi_\alpha(x)$. Then $\pi_\alpha^{-1}(U)$ is an open neighborhood of x, and so $x_n \in \pi_\alpha^{-1}(U)$ for sufficiently large n. Thus $\pi_\alpha(x_n) \in U$ for sufficiently large n, and so $\pi_\alpha(x_n) \to \pi_\alpha(x)$ as needed.

Conversely, by exercise 1.8.11(i) it suffices to prove that for any subbasic neighborhood $\pi_\alpha^{-1}(U_\alpha)$ of x, x_n is eventually in $\pi_\alpha^{-1}(U_\alpha)$. This is equivalent to $\pi_\alpha(x_n)$ being eventually in U_α, which follows from the assumption that $\pi_\alpha(x_n) \to \pi_\alpha(x)$ for all $\alpha \in A$. Therefore $x_n \to x$.

Exercise 1.8.21. (i) Let (X, d) be a metric space, and let $x \in X$. Then the collection of balls $B(x, r)$ with rational radii $r \in \mathbf{Q}$ forms a countable neighborhood base at x — indeed, every ball of real radius centered at x contains such a ball of rational radius.

(ii) Suppose X is second-countable, and let $\mathcal{B}$ be a countable base. Let $\mathcal{B}_x$ be the subset of $\mathcal{B}$ consisting of sets that contain x. This set is countable, and if $x \in U$, then U is a union of elements of $\mathcal{B}$, with at least one element $B \in \mathcal{B}_x$; it follows that $x \in B \subset U$ as needed.

(iii) Let $\{x_1, x_2, \dots\} \subset X$ be a countable dense set. We claim that the countable set $\{B(x_i, q) : i \ge 1, q \in \mathbf{Q}^{>0}\}$ is in fact a base for X. It suffices to prove that, given an open ball $B(x, r)$ with $x \in X$ and $r \in \mathbf{R}^{>0}$, there exists an open ball $x \in B(x_i, q) \subset B(x, r)$. This follows from choosing x_i with $d(x_i, x) < r/2$, and $q \in \mathbf{Q}^{>0}$ with $d(x_i, x) < q < r/2$.

(iv) Let X be a second-countable space with countable base $\mathcal{B} = \{B_1, B_2, \dots\}$. Choosing a point x_i from each B_i, we obtain a set $\{x_1, x_2, \dots\} \subset X$. We claim that this set is dense. Suppose otherwise. Then there exist $x \in X$ and an open neighborhood U of x such that $U \cap \{x_1, x_2, \dots\} = \varnothing$. But U is the union of base elements B_i, and so $x_i \in B_i \subset U$ for some i, a contradiction.

(v) This is false. Consider the constant net $(x)_{\alpha\in\omega_1}$, where ω_1 is the first uncountable ordinal. Then any function $\phi\colon \mathbf{N} \to \omega_1$ is not cofinal, since $\bigcup_{n=1}^{\infty} \phi(n) + 1$ is a strict upper bound.

(vi) [Had to look up a hint for this...] Let X be compact and first-countable, and let $(x_n)_{n\in\mathbf{N}}$ be a sequence in X. Suppose for contradiction that $(x_n)_{n\in\mathbf{N}}$ has no convergent subsequence. Then there are no cluster points[6] of $(x_n)_{n\in\mathbf{N}}$ — if x is a cluster point of $(x_n)_{n\in\mathbf{N}}$, then, given a local base $\mathcal{B}_x = \{B_1 \supset B_2 \supset \dots\}$ at x, we can choose a subsequence $(x_{n_i})_{i\in\mathbf{N}}$ of $(x_n)_{n\in\mathbf{N}}$ satisfying $x_{n_i} \in B_i$; clearly this subsequence converges to x. Thus every point $x \in X$ has a neighborhood U_x that contains only finitely many points of the sequence $(x_n)_{n\in\mathbf{N}}$. Since $(U_x)_{x\in X}$ is an open cover, compactness gives us a finite subcover $X = U_{x'_1} \cup \cdots \cup U_{x'_k}$. It follows that X intersects finitely many points of $(x_n)_{n\in\mathbf{N}}$, which is absurd.

[6] A point $x \in X$ is a *cluster point of a sequence* $(x_n)_{n\in\mathbf{N}}$ if every neighborhood U of x contains infinitely many points of $(x_n)_{n\in\mathbf{N}}$. (Confusingly, there seems to be a different meaning for the cluster point of a set.)

Exercise 1.8.22. As in exercise 1.6.15, the collection of sets of the form $\alpha^{\uparrow} := \{x_\beta \in X : \beta \geq \alpha\}$ satisfies the finite intersection property (FIP). By compactness, we obtain an element $x \in \bigcap_{\alpha\in A} \overline{\alpha^{\uparrow}}$. Since $(x_\alpha)_{\alpha\in A}$ is universal, $(f(x_\alpha))_{\alpha\in A}$ converges to some limit $L \in \{0,1\}$. We claim that $L = f(x)$. There exists $\alpha' \in A$ such that $f(x_\beta) = L$ whenever $\beta \geq \alpha'$. Thus $f(\alpha'^{\uparrow}) = \{L\}$. Since $x \in \bigcap_{\alpha\in A} \overline{\alpha^{\uparrow}} \subset \overline{\alpha'^{\uparrow}}$, we have $f(x) \in f(\overline{\alpha'^{\uparrow}}) = \{L\}$ as needed... [Don't know how to continue.]

Exercise 1.8.23.

Exercise 1.8.24.

Exercise 1.8.25.

Exercise 1.8.26. First notice that d takes values in $[0, +\infty)$ because we are considering bounded functions. It is easy to verify that d is a metric. Suppose Y is a complete metric space, and consider a Cauchy sequence $f_1, f_2, \dots$ in $\mathrm{BC}(X \to Y)$. Then

$$d_Y(f_m(x'), f_n(x')) \leq \sup_{x\in X} d_Y(f_m(x), f_n(x)) =: d(f_m, f_n)$$

whenever $x' \in X$, and so $(f_n(x'))_{n\in\mathbf{N}}$ is a Cauchy sequence in Y for each fixed $x' \in X$. This in uniform in x' in the sense that, for every $\epsilon > 0$, there exists $N \in \mathbf{N}$ such that $d_Y(f_m(x'), f_n(x')) < \epsilon$ whenever $m, n \geq N$ and $x' \in X$. By completeness, $(f_n(x'))_{n\in\mathbf{N}}$ converges to some limit $f(x') \in Y$; in this way we define a function $f\colon X \to Y$. We claim that (i) $f \in \mathrm{BC}(X \to Y)$, and that (ii) $(f_n)_{n\in\mathbf{N}}$ converges to f in $\mathrm{BC}(X \to Y)$.

(i) To prove f is bounded, it suffices to show that $d(0, f) < \infty$. Since $d(0, f) \leq d(0, f_n) + d(f_n, f)$ for all n, and since Cauchy sequences are bounded, it suffices to find $n \in \mathbf{N}$ such that

$$\sup_{x\in X} d_Y(f_n(x), f(x)) < \infty.$$

We have $d_Y(f_N(x'), f_n(x')) \leq \epsilon$ whenever $n \geq N$ and $x' \in X$. Since the metric $d_Y\colon Y \times Y \to [0, +\infty)$ is continuous, taking $n \to \infty$ implies that $d_Y(f_N(x'), f(x')) \leq \epsilon$ whenever $x' \in X$. Therefore

$$\sup_{x\in X} d_Y(f_N(x), f(x)) \leq \epsilon$$

as needed.

To prove that f is continuous, it suffices to prove that for every $x \in X$ and $\epsilon > 0$, there exists a neighborhood U of x such that $d_Y(f(x'), f(x)) \leq \epsilon$ for all $x' \in U$. Let $x \in X$ and $\epsilon > 0$. Then...

(ii) ... [I did not finish this exercise unfortunately; the main ideas can be found at `https://math.stackexchange.com/a/76455/`.]

Exercise 1.8.27.

Exercise 1.8.28.

Exercise 1.8.29.

Exercise 1.8.30.

1.9. The strong and weak topologies

The strong do what they can
and the weak suffer what they must.

— THUCYDIDES, *History of the Peloponnesian War* 5.89 (c. 400B.C.)

Exercise 1.9.1. It suffices to prove sequential continuity, since the norm induces a metric. If $(v_n, w_n) \to (v, w)$, then $v_n \to v$ and $w_n \to w$. Since $\|v_n + w_n - (v + w)\| \leq \|v_n - v\| + \|w_n - w\|$, it follows that $v_n + w_n \to v + w$. Similarly, if $(c_n, v_n) \to (c, v)$, then $c_n \to c$ and $v_n \to v$; since

$$\|c_n v_n - cv\| \leq \|c_n v_n - c_n v\| + \|c_n v - cv\| = |c_n|\|v_n - v\| + |c_n - c|\|v\|,$$

we have $c_n v_n \to cv$. The same argument works with quasi-norms.

Exercise 1.9.2. The argument from exercise 1.9.1 works the same (all we need is the triangle inequality and homogeneity) to prove that semi-normed vector spaces are topological vector spaces. If $v \neq 0$ with $\|v\| = 0$, then any open ball $B(0, \epsilon) := \{w \in V : \|w\| < \epsilon\}$ contains v, and similarly we have $0 \in B(v, \epsilon)$; thus 0 and v cannot be separated by disjoint open sets and so V is not Hausdorff. Conversely, if the semi-norm is a norm, then given $v \neq w$ in V, we may let $\epsilon := \|v - w\| > 0$; it follows that $B(v, \epsilon/2)$ and $B(w, \epsilon/2)$ are disjoint. Thus V is Hausdorff.

Exercise 1.9.3. If $U \in \bigcup_{\alpha \in A} \mathcal{F}_\alpha$, then $U \in \mathcal{F}_\alpha$ for some $\alpha \in A$, and so $+_\alpha^{-1} \colon (V, \mathcal{F}_\alpha) \times (V, \mathcal{F}_\alpha) \to (V, \mathcal{F}_\alpha)$ is such that $+_\alpha^{-1}(U)$ is open; it follows that $+^{-1} \colon (V, \mathcal{F}) \times (V, \mathcal{F}) \to (V, \mathcal{F})$ is such that $+^{-1}(U)$ is open as well. If $U_1, U_2 \in \mathcal{F}$ are such that $+^{-1}(U_1)$ and $+^{-1}(U_2)$ are open, then $+^{-1}(U_1 \cap U_2) = +^{-1}(U_1) \cap +^{-1}(U_2)$ is open as well. Similarly, if $U_1, U_2, \cdots \in \mathcal{F}$ are such that $+^{-1}(U_n)$ is open for each n, then $+^{-1}(\bigcup_{n=1}^\infty U_n) = \bigcup_{n=1}^\infty +^{-1}(U_n)$ is open as well. Therefore the collection of open subsets of $\mathcal{F}$ whose preimage under $+$ is open contains $\bigcup_{\alpha \in A} U_\alpha$ and is itself a topology; thus it contains $\mathcal{F}$ as needed. The argument for continuity of scalar multiplication is similar.

If $x_n \to x$ in $\mathcal{F}$, then $x_n \to x$ in $\mathcal{F}_\alpha$ since $\mathcal{F}$ is stronger than $\mathcal{F}_\alpha$ (i.e., $\mathcal{F} \supset \mathcal{F}_\alpha$). Conversely, if $x_n \to x$ in $\mathcal{F}_\alpha$ for all $\alpha \in A$, then we may consider the collection of open neighborhoods of x that eventually contain x_n — this collection contains $\bigcup_{\alpha \in A} U_\alpha$ by hypothesis. It is closed under finite intersections — if $U_1, \ldots, U_k$ contain x_n past $n_1, \ldots n_k$, then $U_1 \cap \cdots \cap U_k$ contains x_n past $\max\{n_1, \ldots, n_k\}$. Finally, it is clearly closed under arbitrary unions — if x_n is eventually in U_β for $\beta \in B$, then x_n is eventually in $U_\beta \subset \bigcup_{\beta \in B} U_\beta$. Thus this collection is a topology on V containing $\bigcup_{\alpha \in A} U_\alpha$, which means that it contains $\mathcal{F} := \bigvee_{\alpha \in A} \mathcal{F}_\alpha$ by definition of $\mathcal{F}$ (recall that it is the *weakest* topology that contains $\bigcup_{\alpha \in A} \mathcal{F}_\alpha$).

Exercise 1.9.4. Write $\mathcal{F}_{W_\beta}$ for the topology generated by the seminorm $\|\cdot\|_{W_\beta}$. It suffices to prove that $T^{-1}(U) \subset V$ is open with respect to $(\|\cdot\|_{V_\alpha})_{\alpha \in A}$ whenever $U \in \bigcup_{\beta \in B} \mathcal{F}_{W_\beta}$, since $\bigcup_{\beta \in B} \mathcal{F}_{W_\beta}$ is a subbase

It seems that people sometimes denote seminorms by p and reserve $\|\cdot\|$ for actual norms.

for $\mathcal{F}_W$ by definition. Say $U \in \mathcal{F}_{W_\beta}$. Then, given $f \in T^{-1}(U)$, we have $B_{W_\beta}(Tf, \epsilon) := \{w \in W : \|w - Tf\|_{W_\beta} < \epsilon\} \subset U$ for some $\epsilon > 0$. By hypothesis, there exists a finite set $A_\beta \subset A$ and a constant C_β such that $\|Tg\|_{W_\beta} \le C_\beta \sum_{\alpha \in A_\beta} \|g\|_{V_\alpha}$ for all $g \in V$. Setting $\delta := \epsilon/(C_\beta |A_\beta|)$, we have $\|Tg\|_{W_\beta} < \epsilon$ whenever $\|g\|_{V_\alpha} < \delta$ for all $\alpha \in A_\beta$. We see that $\bigcap_{\alpha \in A_\beta} B_{V_\alpha}(f, \delta)$ is a finite intersection of sets that are open with respect to $(\|\cdot\|_{V_\alpha})_{\alpha \in A}$, and is thus itself open with respect to $(\|\cdot\|_{V_\alpha})_{\alpha \in A}$; by previous arguments, we have $T(\bigcap_{\alpha \in A_\beta} B_{V_\alpha}(f, \delta)) \subset B_{W_\beta}(Tf, \epsilon) \subset U$. Therefore T is continuous.

Conversely, suppose T is continuous. Let $\beta \in B$, and consider $T^{-1}(B_{W_\beta}(0,1)) \subset V$. This set is open with respect to $(\|\cdot\|_{V_\alpha})_{\alpha \in A}$, and so there exists a finite set $A_\beta \subset A$ such that

$$0 \in \bigcap_{\alpha \in A_\beta} B_{V_\alpha}(f_\alpha, \delta_\alpha) \subset T^{-1}(B_{W_\beta}(0,1)).$$

We may choose the f_α to all be equal to 0, making the δ_α smaller if needed. Let $\delta := \min_{\alpha \in A_\beta} \delta_\alpha$. Given $f \in V$ with $\max_{\alpha \in A_\beta} \|f\|_{V_\alpha} < \delta$, we see that $f \in \bigcap_{\alpha \in A_\beta} B_{V_\alpha}(0, \delta)$, so that $\|Tf\|_{W_\beta} < 1$. In general, since

$$\left\| \frac{f}{2\delta^{-1} \max_{\alpha \in A_\beta} \|f\|_{V_\alpha}} \right\|_{V_{\alpha'}} \le \frac{\delta}{2} < \delta$$

whenever $\alpha' \in A_\beta$, we deduce from homogeneity that

$$\|Tf\|_{W_\beta} < \frac{2}{\delta} \max_{\alpha \in A_\beta} \|f\|_{V_\alpha} \le \frac{2}{\delta} \sum_{\alpha \in A_\beta} \|f\|_{V_\alpha}$$

as needed.

Exercise 1.9.5.

$$\begin{tikzcd}
\prod_\alpha V_\alpha \times \prod_\alpha V_\alpha \arrow[r, "+"] \arrow[d, "\pi_\beta \times \pi_\beta"] & \prod_\alpha V_\alpha \arrow[d, "\pi_\beta"] \\
V_\beta \times V_\beta \arrow[r, "+_\beta"] & V_\beta
\end{tikzcd}$$

We only prove the continuity of addition in the product space. It suffices to verify that the openness of preimages of subbasic open sets of the form $\pi_\beta^{-1}(U_\beta)$ with $U_\beta \subset V_\beta$ open. Since $((v_\alpha)_{\alpha \in A}, (v'_\alpha)_{\alpha \in A}) \in +^{-1}(\pi_\beta^{-1}(U_\beta))$ if and only if $(v_\beta, v'_\beta) \in +_\beta^{-1}(U_\beta)$, we see that

$$+^{-1}(\pi_\beta^{-1}(U_\beta)) = (\pi_\beta^{-1} \times \pi_\beta^{-1})(+_\beta^{-1}(U_\beta)).$$

Since V_β is a TVS, the map $+_\beta$ is continuous. By definition of product spaces, $\pi_\beta \times \pi_\beta$ is continuous, with inverse given by $\pi_\beta^{-1} \times \pi_\beta^{-1}$. Thus, the preimage is open as needed.

Exercise 1.9.6. [W. Rudin, *Functional Analysis*, Theorem 1.10] We first prove a lemma: *given a neighborhood W of $0 \in V$, there exists a neighborhood U of 0 such that $U + U \subset W$.* Indeed, since $+_V^{-1}(W) \subset V \times V$ is open, we have $(0,0) \in V_1 \times V_2 \subset +_V^{-1}(W)$ for some open sets $V_1, V_2 \subset V$. Letting $U := V_1 \cap V_2$, we see that U is a neighborhood of 0 such that $U + U \subset W$.

Suppose V is a T_1 topological vector space,[7] so that singleton sets $\{x\}$ are closed, and let $x \in V \setminus \{0\}$. We will find open sets separating 0 and x. Applying the lemma twice with $W = V \setminus \{x\}$, we obtain a neighborhood U of 0 satisfying $U + U + U + U \subset W$, so that $x \notin U + U + U$. It follows that $(U + U) \cap (x - U) = \varnothing$ are our desired separating sets.

[7] Some authors, including Rudin but not Tao, take it as an axiom that topological vector spaces are T_1.

Exercise 1.9.7. We first verify that the sets $B(f, \epsilon, r)$ form a base. Since $f \in B(f, \epsilon, r)$, they cover $L(X)$. Suppose $g \in B(f, \epsilon, r) \cap B(f', \epsilon', r')$. Then, we have

$$\mu\{|f - g| \geq r\} < \epsilon \quad \text{and} \quad \mu\{|f' - g| \geq r'\} < \epsilon'.$$

Choose $\epsilon'' > 0$ such that

$$\mu\{|f - g| \geq r\} < \epsilon - \epsilon'' \quad \text{and} \quad \mu\{|f' - g| \geq r'\} < \epsilon' - \epsilon''.$$

We claim there exists $r'' > 0$ such that

$$\mu\{|f - g| \geq r - r''\} < \epsilon - \epsilon'' \quad \text{and} \quad \mu\{|f' - g| \geq r - r''\} < \epsilon - \epsilon''.$$

Indeed, it suffices to prove that $F(r) := \mu\{f \geq r\}$ is left-continuous in r whenever f is a measurable function. Since $F\colon \mathbf{R} \to \mathbf{R}$ is a real function, it suffices to verify sequential left-continuity. Let $(r_n)_{n \in \mathbf{N}}$ be an increasing sequence of reals such that $r_n < r$ and $r_n \to r$. Then

$$\begin{aligned}
\lim_{n \to \infty} F(r_n) &= \lim_{n \to \infty} \mu\{f \geq r_n\} \\
&= \mu\Big(\bigcap_{n \in \mathbf{N}} \{f \geq r_n\} \Big) \\
&= \mu\{f \geq r\} =: F(r)
\end{aligned}$$

by dominated convergence for sets, which holds since $\mu(X) < \infty$.

Now we may prove that $B(g, \epsilon'', r'') \subset B(f, \epsilon, r) \cap B(f', \epsilon', r')$. Let $h \in B(g, \epsilon'', r'')$. Then

$$\begin{aligned}
\mu\{|f - h| \geq r\} &\leq \mu\{|f - g| \geq r - r''\} + \mu\{|g - h| \geq r''\} \\
&< (\epsilon - \epsilon'') + \epsilon'' \\
&= \epsilon
\end{aligned}$$

by the triangle inequality. Similarly, we have $\mu\{|f' - h| \geq r'\} < \epsilon'$, and the result follows.

Now we prove that this base generates a topology for $L(X)$ that turns $L(X)$ into a topological vector space. To prove that addition $+\colon L(X) \times L(X) \to L(X)$ is continuous, given $g, h \in L(X)$ such that $g + h \in B(f, \epsilon, r)$, we must find $\epsilon', r', \epsilon'', r''$ such that

$$B(g, \epsilon', r') + B(h, \epsilon'', r'') \subset B(f, \epsilon, r).$$

By hypothesis, we have

$$\mu\{|g + h - f| \geq r\} < \epsilon,$$

and we may choose $r', \epsilon' > 0$ as before, so that

$$\mu\{|g + h - f| \geq r - r'\} < \epsilon - \epsilon'.$$

Then, given

$$g' \in B\Big(g, \frac{\epsilon'}{2}, \frac{r'}{2}\Big) \quad \text{and} \quad h' \in B\Big(h, \frac{\epsilon'}{2}, \frac{r'}{2}\Big),$$

we compute

$$\begin{aligned}
&\mu\{|g'+h'-f| \geq r\} \\
&\leq \mu\{|g+h-f| \geq r-r'\} + \mu\Big\{|g'-g| \geq \frac{r'}{2}\Big\} + \mu\Big\{|h'-h| \geq \frac{r'}{2}\Big\} \\
&< (\epsilon - \epsilon') + \frac{\epsilon'}{2} + \frac{\epsilon'}{2} \\
&= \epsilon
\end{aligned}$$

as needed.

To prove continuity of scalar multiplication $\cdot\colon \mathbf{C} \times L(X) \to L(X)$, given $z \in \mathbf{C}$ and $g \in L(X)$ such that $zg \in B(f, \epsilon, r)$, we must find $\epsilon', r', \epsilon'' > 0$ such that

$$B(z, \epsilon') \cdot B(g, \epsilon'', r') \subset B(f, \epsilon, r).$$

By hypothesis, we have $\mu\{|zg - f| \geq r\} < \epsilon$. We choose $\epsilon', M, r' > 0$, in that order, such that $\epsilon', r' < 1$, $\mu\{|g| \geq M\} \leq \epsilon'/2$,[8] and

$$\mu\{|zg - f| \geq r - (|z| + 1 + \epsilon' M)r'\} < \epsilon - \epsilon'.$$

[8] Since $\mu(X) < \infty$, we apply dominated convergence for sets to the sequence

$$\cdots \supset \{|g| \geq M\} \supset \{|g| \geq M+1\} \supset \cdots$$

to obtain

$$\begin{aligned}
\lim_{M\to\infty} \mu\{|g| \geq M\} &= \mu\Big(\bigcap_{M \geq 1} \{|g| \geq M\}\Big) \\
&= 0.
\end{aligned}$$

Given

$$z' \in B(z, \epsilon' r') \quad \text{and} \quad g' \in B(g, \epsilon'/2, r'),$$

we compute

$$\begin{aligned}
\mu\{|z'g' - f| \geq r\} &\leq \mu\{|zg - f| \geq r - (|z| + 1 + \epsilon' M)r'\} \\
&\qquad + \mu\{|z'||g'-g| \geq (|z|+1)r'\} \\
&\qquad + \mu\{|z'-z||g| \geq \epsilon' M r'\} \\
&\leq (\epsilon - \epsilon') \\
&\qquad + \mu\{|g'-g| \geq r'\} \\
&\qquad + \mu\{|g| \geq M\} \\
&< (\epsilon - \epsilon') + \epsilon'/2 + \epsilon'/2 \\
&= \epsilon
\end{aligned}$$

as needed.

Finally, a sequence $f_n \in L(X)$ converges to a limit f in this topology iff it converges in measure. Indeed, this equivalence is immediate from the definitions: recall that f_n is said to *converge in measure to f* if, for every $r > 0$, we have $\lim_{n\to\infty} \mu\{|f_n - f| \geq r\} = 0$.

Exercise 1.9.8. Consider the typewriter sequence $\mathbf{1}_{[\frac{n-2^k}{2^k}, \frac{n-2^k+1}{2^k}]}$, where $k \geq 0$ and $2^k \leq n < 2^{k+1}$. This sequence does not converge pointwise a.e. to zero, since every $x \in [0,1]$ is contained in infinitely many sets of the form $[\frac{n-2^k}{2^k}, \frac{n-2^k+1}{2^k}]$. However, given a subsequence, we observe that for each k, there can be anywhere from 0 to 2^k terms of the form $\mathbf{1}_{[\frac{n-2^k}{2^k}, \frac{n-2^k+1}{2^k}]}$. Choose a subsequence $(1_{E_n})_{n\in\mathbf{N}}$ of this subsequence such that, for each k, there is at most one term of this form. Then $\sum_{n=1}^{\infty} \mu(E_n) \leq \sum_{k=0}^{\infty} \frac{1}{2^k} = 2 < \infty$, and the Borel–Cantelli lemma implies that almost every $x \in [0,1]$ belongs to finitely many E_n. Thus, this subsubsequence converges pointwise a.e. to zero. If pointwise a.e. convergence were topologizable, then the Urysohn subsequence principle would imply that the typewriter sequence converges pointwise a.e. to zero. Therefore pointwise a.e. convergence is not topologizable, and we are done.

Exercise 1.9.9.

Exercise 1.9.10.

Exercise 1.9.11.

Exercise 1.9.12.

Exercise 1.9.13. *The weak topology on V makes V a TVS.* We first verify that addition is continuous. If $x + y = z$ and $\epsilon > 0$, then

$$B_\lambda(x, \epsilon/2) + B_\lambda(y, \epsilon/2) \subset B_\lambda(z, \epsilon).$$

Indeed, given x', y' satisfying $\|x' - x\|_\lambda < \epsilon/2$ and $\|y' - y\|_\lambda < \epsilon/2$, we have $\|x' + y' - (x + y)\|_\lambda < \epsilon$ by the triangle inequality. Similarly, if $cx = y$ and $\epsilon > 0$, then, choosing $0 < \epsilon' < \epsilon$ such that $\epsilon' < 2\|x\|_\lambda$, we have

$$B\Big(c, \frac{\epsilon'}{2\|x\|_\lambda}\Big) \cdot B_\lambda\Big(x, \frac{\epsilon'}{2(|c|+1)}\Big) \subset B_\lambda(y, \epsilon)$$

since

$$\begin{aligned} \|c'x' - cx\|_\lambda &\leq |c' - c|\|x\|_\lambda + |c'|\|x' - x\|_\lambda \\ &< \frac{\epsilon}{2\|x\|_\lambda}\|x\|_\lambda + (|c|+1)\frac{\epsilon}{2(|c|+1)} \\ &= \epsilon. \end{aligned}$$

The general recipe for checking that a collection of seminorms $(\|\cdot\|_\alpha)_{\alpha \in A}$ generates a topology for a TVS V is as follows: since the balls $B_\alpha(x, r) := \{v \in V : \|v - x\|_\alpha < r\}$ with $r > 0$ and $x \in V$ form a subbase for this topology, it suffices to prove that the preimages of such sets under addition and scalar multiplication are open. For addition, if we have $(y, z) \in +^{-1}(B_\alpha(x, r))$, we must find $B_{\alpha'}(y, r')$ and $B_{\alpha''}(z, r'')$ such that $B_{\alpha'}(y, r') + B_{\alpha''}(z, r'') \subset B_\alpha(x, r)$.

The weak topology is weaker than the strong topology on V. It suffices to verify that every set of the form $\{v \in V : \|v - x\|_\lambda < \epsilon\}$ is open with respect to the norm $\|\cdot\|_V$. Say $\|v - x\|_\lambda < \epsilon$. We must find $\delta > 0$ such that $\|v' - v\|_V < \delta$ implies $\|v' - x\|_\lambda < \epsilon$. Since

$$\|v' - x\|_\lambda \leq \|v' - v\|_\lambda + \|v - x\|_\lambda$$

and

$$\|v' - v\|_\lambda := |\lambda(v' - v)| \leq \|\lambda\|_{\mathrm{op}}\|v' - v\|_V,$$

choosing δ such that $\|v - x\|_\lambda < \epsilon - \delta\|\lambda\|_{\mathrm{op}}$ gives the claim.

The weak topology on V^* makes V^* a TVS.* The proof is essentially the same as (i).

The weak topology on V^* is weaker than the weak topology on V^*.* Recall that the weak topology on V^* is generated by the seminorms $\|\lambda\|_a$ for all $a \in (V^*)^*$. We must prove that every weak* ball $B_x(\lambda, \epsilon)$ is open with respect to the weak topology on V^*; that is, given $\omega \in B_x(\lambda, \epsilon)$, we must find $a \in (V^*)^*$ and $\delta > 0$ such that $\omega \in B_a(\omega, \delta) \subset B_x(\lambda, \epsilon)$.

Here we use the canonical embedding $V \hookrightarrow (V^*)^*$ sending v to the evaluation map $\hat{v}$ defined by $\hat{v}(\omega) := \omega(v)$. Then, we let $a := \hat{x}$ and we choose sufficiently small δ satisfying $\|\omega - \lambda\|_x < \epsilon - \delta$. Thus $\|\omega' - \omega\|_{\hat{x}} < \delta$ implies that

$$\begin{aligned} \|\omega' - \lambda\|_x &\leq \|\omega' - \omega\|_x + \|\omega - \lambda\|_x \\ &= \|\omega' - \omega\|_{\hat{x}} + \|\omega - \lambda\|_x \\ &< \delta + (\epsilon - \delta) \\ &= \epsilon \end{aligned}$$

as needed.

When V is reflexive, the weak and weak topologies on V^* are equivalent.* By the previous part, it suffices to prove that any ball $B_a(\lambda, \epsilon)$ in the weak topology on V^* is open with respect to the weak* topology on V^*. The proof is essentially the same, since now every $a \in (V^*)^*$ is of the form $a = \hat{x}$, so for $\omega \in B_{\hat{x}}(\lambda, \epsilon)$ we have the basic open neighborhood $\omega \in B_x(\omega, \delta) \subset B_{\hat{x}}(\lambda, \epsilon)$ as $\|\cdot\|_x = \|\cdot\|_{\hat{x}}$.

Exercise 1.9.14. We prove that the weak topology on a normed vector space V is Hausdorff. It suffices to show that, given $x \in V \setminus \{0\}$, there exist open sets separating 0 and x. By the Hahn–Banach theorem, there exists $\lambda \in V^*$ such that $\lambda x = 1$. Then the balls $B_\lambda(x, 1/2)$ and $B_\lambda(0, 1/2)$ are disjoint, since if $|\lambda x - \lambda v| < 1/2$ and $|\lambda v| < 1/2$, then the triangle inequality yields $|\lambda x| < 1$, a contradiction.

We can similarly show that the weak* topology on V^* is Hausdorff. Given $\lambda \in V^* \setminus \{0\}$, since λ is non-zero, there exists $x \in V$ such that $\lambda x = 1$. Then the balls $B_x(0, 1/2)$ and $B_x(\lambda, 1/2)$ are disjoint.

Exercise 1.9.15. (i) Recall that elements of $V^* \equiv \ell^1(\mathbf{N})$ are absolutely summable sequences $(a_n)_{n\in\mathbf{N}}$ of complex numbers, and they act on elements $(b_n)_{n\in\mathbf{N}} \in V = c_0(\mathbf{N})$ by sending them to $\sum_{n=1}^\infty a_n b_n$. In this way, $(a_n)_{n\in\mathbf{N}} \in V^*$ sends $e_n \in V$ to a_n; by absolute summability we see that $a_n \to 0$, and so $e_n \rightharpoonup 0$ in V. As for strong convergence, we see that $\|e_m - e_n\|_{\ell^\infty} = 1$ whenever $m \neq n$, which implies that $(e_n)_{n\in\mathbf{N}}$ is not Cauchy and thus not strongly convergent in V.

Perhaps it would be clearer to write $e^{(m)}$ or $([n = m])_{n\in\mathbf{N}}$ instead of e_m.

(ii) As before, $e_m \in V^*$ acts on $x := (a_n)_{n\in\mathbf{N}} \in V$ by sending it to a_m, and so $e_m(x) = a_m \to 0$ as $x \in V = c_0(\mathbf{N})$; that is, $e_n \xrightarrow{w^*} 0$ as needed. To prove that e_n does not converge in the weak or strong senses in V^*, it suffices by exercise 1.9.13 to show that it does not converge in the weak sense. If it did, then, there would exist some limit $(l_n)_{n\in\mathbf{N}} \in c_0(\mathbf{N})$ such that, given any $a = (a_n)_{n\in\mathbf{N}} \in (V^*)^* \equiv \ell^\infty(\mathbf{N})$, we have $a_n = a(e_n) \to a((l_n)_{n\in\mathbf{N}}) = \sum_{n=1}^\infty a_n l_n$. Plugging in $a = e_m$, we see that $l_m = 0$ for each m, and so $(l_n)_{n\in\mathbf{N}}$ must be the zero sequence $(0)_{n\in\mathbf{N}}$. But this implies that $a_n \to 0$ for every bounded sequence $(a_n)_{n\in\mathbf{N}}$, which is absurd.

(iii) To see that $(\sum_{m=n}^\infty e_m)_{n\in\mathbf{N}} \xrightarrow{w^*} 0$ in $\ell^\infty(\mathbf{N})$, we observe that $(a_n)_{n\in\mathbf{N}} \in \ell^1(\mathbf{N})$ implies that $|\sum_{m=n}^\infty a_n| \leq \sum_{m=n}^\infty |a_n| \to 0$ as $n \to \infty$. This sequence however does not converge in the weak sense. Indeed, suppose for contradiction that it did converge weakly to a limit $(l_n)_{n\in\mathbf{N}} \in \ell^\infty(\mathbf{N})$. By considering the maps $\hat{e}_m \in (\ell^\infty(\mathbf{N}))^*$ defined by sending $(a_n)_{n\in\mathbf{N}}$ to a_m, we see for fixed m that

$$[m \geq n] = \hat{e}_m\Big(\sum_{k=n}^{\infty} e_k\Big) \to \hat{e}_m((l_n)_{n\in\mathbf{N}}) = l_m \quad \text{as } n \to \infty,$$

which implies that $(l_n)_{n\in\mathbf{N}}$ is the zero sequence $(0)_{n\in\mathbf{N}}$. On the other hand, if we consider a generalized limit functional $\lambda\colon \ell^\infty(\mathbf{N}) \to \mathbf{C}$, we have $\lambda(\sum_{m=n}^{\infty} e_m) = 1$ for all n, which implies that

$$0 = \lim_{n\to\infty} l_n = \lambda((l_n)_{n\in\mathbf{N}}) = \lim_{n\to\infty} \lambda\Big(\sum_{m=n}^{\infty} e_m\Big) = 1,$$

a contradiction.

Exercise 1.9.16. Suppose E is strongly bounded, so that there exists C such that $\|x\|_V \leq C$ whenever $x \in E$. Then, given $\lambda \in V^*$, we have

$$|\lambda(x)| \leq \|\lambda\|_{\text{op}}\|x\|_V \leq C\|\lambda\|_{\text{op}}$$

for all $x \in E$, as needed.

Conversely, suppose E is weakly bounded. By the Hahn–Banach theorem, the evaluation map $\iota\colon V \to V^{**}$ is an isometry. The collection $(\iota(x))_{x\in E}$ is pointwise bounded, since $\{\iota(x)(\lambda) : x \in E\} = \lambda(E)$ is bounded by hypothesis. Thus, the uniform boundedness principle implies that $\{\|\iota(x)\|_{V^{**}} : x \in E\}$ is bounded. Since ι is an isometry, we conclude that $\{\|x\|_V : x \in E\}$ is bounded, which gives the claim.

Similarly, if V is a Banach space, we may show that a subset $F \subset V^*$ is strongly bounded iff it is weak* bounded, in the sense that $\{\lambda(x) : \lambda \in F\}$ is bounded for each $x \in V$. The proof is essentially the same, except we now apply the uniform boundedness principle to the collection $(\lambda)_{\lambda\in F}$.

From this discussion, we see that weak and weak* convergence implies boundedness.

Exercise 1.9.17. By exercise 1.5.14, there exists $\lambda \in V^*$ such that $\|\lambda\|_{V^*} = 1$ and $\lambda x = \|x\|_V$. It follows that $\lambda x_n \to \|x\|_V$ and $|\lambda x_n| \leq \|\lambda\|_{V^*}\|x_n\|_V = \|x_n\|_V$; taking the limit inferior on both sides of the inequality gives the claim.

[To do: show the result for $\lambda_n \in V^*$ and construct an example of strict inequality.]

Exercise 1.9.18. If $x_n \to x$ in a Hilbert space H, then $\|x_n\| \to \|x\|$ by the triangle inequality: $|\|x_n\| - \|x\|| \leq \|x_n - x\|$. Conversely, suppose $\|x_n\| \to \|x\|$ and $x_n \rightharpoonup x$; we prove that $x_n \to x$. It suffices to show that $\langle x_n - x, x_n - x\rangle \to 0$. By weak convergence, we compute

$$\begin{aligned}\langle x_n - x, x_n - x\rangle &= \|x_n\|^2 + \|x\|^2 - (\langle x, x_n\rangle + \langle x_n, x\rangle)\\ &\to 2\|x\|^2 - 2\|x\|^2 = 0,\end{aligned}$$

which gives the claim.

Exercise 1.9.19. (i)

(ii)

(iv)

Exercise 1.9.20. We prove that the closed unit ball[9] $\overline{B}(0,1) := \{x \in V : \|x\|_V \leq 1\}$ is closed in the weak topology. Let $x \in V$ be an adherent point of $\overline{B}(0,1)$ in the weak topology, and suppose for contradiction that $\|x\|_V > 1$. Choose $\epsilon > 0$ such that $\|x\|_V > 1+\epsilon$. By exercise 1.5.14, there exists $\lambda \in V^*$ such that $\|\lambda\|_{\text{op}} = 1$ and $\lambda x = \|x\|_V$. Since x is an adherent point of $\overline{B}(0,1)$, there exists $x' \in B_\lambda(x,\epsilon) \cap B(0,1)$. We may then compute

$$\begin{aligned}\|x\|_V &= |\lambda x| \\ &\leq |\lambda x'| + |\lambda(x - x')| \\ &< \|\lambda\|_{\text{op}}\|x'\|_V + \epsilon \\ &\leq 1+\epsilon,\end{aligned}$$

which contradicts our choice of ϵ.

We may prove that the closed unit ball in V^* is closed in the weak* topology in a very similar fashion — the main fact we use is that if $\|\lambda\|_{V^*} > 1+\epsilon$, then there exists $x \in V$ such that $\|x\|_V = 1$ and $|\lambda x| > 1+\epsilon$.

[9] In normed vector spaces, the closure of the open unit ball is the closed unit ball (thankfully). Also, just a warning: it is possible for a point x to have all subbasic open neighborhoods intersect a set S, while not be an adherent point of S; see exercise 1.8.11(ii).

Exercise 1.9.21. [Sketch] Suppose V is a Banach space, and let (λ_n) be a Cauchy sequence in the weak* topology on V^*. Then $(\lambda_n x)$ is Cauchy for all $x \in V$ (since $|\lambda_m x - \lambda_n x| = \|\lambda_m - \lambda_n\|_x$), which defines a map $\lambda \colon V \to \mathbf{C}$ by completeness of $\mathbf{C}$. The uniform boundedness principle then implies that $\lambda \in V^*$, and so $\lambda_n \to \lambda$ as needed.

Exercise 1.9.22. (i) Let $x \in V \setminus W$, and define $\lambda \colon W \oplus \operatorname{span}\{x\} \to \mathbf{C}$ by sending W to 0 and x to 1. By the Hahn–Banach theorem, we may extend λ to $\tilde{\lambda} \in V^*$. Then, $B_{\tilde{\lambda}}(x, 1/2) \subset V \setminus W$, since if $|\tilde{\lambda}x' - \lambda x| < 1/2$, then $|\tilde{\lambda}x'| > 1/2$, which implies that $x' \notin W$. Thus $V \setminus W$ is open in the weak topology of V, and so W is closed in the weak topology of V, as needed.

(ii) Suppose $w_n \rightharpoonup w$ in W. Since any $\lambda \in V^*$ can be restricted to $\lambda|_W \in W^*$, with $B_{\lambda|_W}(x,\epsilon) \subset B_\lambda(x,\epsilon)$, we see that w_n is eventually in $B_\lambda(x,\epsilon)$, as needed. Conversely, suppose $w_n \rightharpoonup w$ in V. Since any $\lambda \in W^*$ can be extended to some $\tilde{\lambda} \in V^*$ using the Hahn–Banach theorem, with $B_\lambda(x,\epsilon) \subset B_{\tilde{\lambda}}(x,\epsilon)$, we see that w_n is eventually in $B_{\tilde{\lambda}}(x,\epsilon)$. Since $w_n \in W$, and since $W \cap B_{\tilde{\lambda}}(x,\epsilon) = B_\lambda(x,\epsilon)$, the result follows. Note that we did not need the hypothesis that W is closed.

Exercise 1.9.23. (i) Suppose $x_n \rightharpoonup x$ in $c_0(\mathbf{N})$. Then the x_n are bounded by the uniform boundedness principle (see exercise 1.9.16), and the basis vectors $e_k \in \ell^1(\mathbf{N})$ imply that $x_{n,k} \to x_k$ for all k. Conversely, suppose $\|x_n\|_{\ell^\infty} \leq M$ and $x_{n,k} \to x_k$ for all k. Suppose $\sum_{n=1}^\infty |a_n| < \infty$; we must prove that $\sum_{k=1}^\infty a_k x_{n,k} \to \sum_{k=1}^\infty a_k x_k$. Choose large N such that $\sum_{k \geq N} |a_k| < \epsilon/4M$, and choose large N' such that $|x_{n,k} - x_k| \leq \epsilon/(2\sum_{k<N}|a_k|)$ whenever $n \geq N'$ and $k < N$ (the case where the denominator is zero is trivial). Then, we have

$$\begin{aligned}
\Big|\sum_{k=1}^{\infty} a_k(x_{n,k}-x_k)\Big| &\le \sum_{k=1}^{\infty} |a_k||x_{n,k}-x_k| \\
&= \sum_{k<N} |a_k||x_{n,k}-x_k| + \sum_{k\ge N} |a_k||x_{n,k}-x_k| \\
&\le \frac{\epsilon}{2\sum_{k<N}|a_k|}\sum_{k<N}|a_k| + 2M\sum_{k\ge N}|a_k| \\
&\le \frac{\epsilon}{2} + 2M\cdot\frac{\epsilon}{4M} \\
&= \epsilon
\end{aligned}$$

whenever $n \ge N'$, which completes the proof.

(ii) The idea is the same, just that we now we estimate the sum $\sum_k |\lambda_{n,k}-\lambda_k||x_k|$, estimating $\sum_{k<N}$ by using the convergence $\lambda_{n,k} \to \lambda_k$ for $k < N$ and $\sum_{k\ge N}$ by choosing large N with $\sum_{k\ge N}|\lambda_k|$ small using the boundedness of x_k.

(iii) Consider the sequence $x_n = ([k \le n])_{k\in\mathbf{N}}$ in $c_0(\mathbf{N})$:

$$\begin{matrix}
1 & 0 & 0 & 0 & \cdots \\
1 & 1 & 0 & 0 & \cdots \\
1 & 1 & 1 & 0 & \cdots \\
1 & 1 & 1 & 1 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{matrix}$$

The sequence x_n has no weak limit in $c_0(\mathbf{N})$, since $(1)_{n\in\mathbf{N}} \notin c_0(\mathbf{N})$. We prove that x_n is Cauchy. It suffices to show that, for any $\epsilon > 0$ and $\sum_n |a_n| < \infty$, there exists N such that $\sum_k |a_k||x_{m,k}-x_{n,k}| < \epsilon$ whenever $m, n \ge N$. Since

$$\sum_k |a_k||x_{m,k}-x_{n,k}| = \sum_{n<k\le m} |a_k| \le \sum_{k>n} |a_k|$$

for $m > n$, we see that choosing large N for which $\sum_{k>N}|a_k| < \epsilon$ gives the claim.

Remark (Details in the proof of the Banach–Alaoglu theorem). *One easily verifies that the weak* topology on B^* is nothing more than the product topology of D^B restricted to B^*.* The idea here is we have the map $i\colon B^* \to D^B$ defined by $i(\phi) := \phi|_B$, and we want to prove that this map is a homeomorphism onto its image. We first prove that i is injective. This follows from the fact that $\phi \in V^*$ is completely determined by $\phi|_B$.

Now we must prove that i is a homeomorphism onto its image. The product space D^B has the standard product subbase, which consists of sets of the form $\{\phi \in D^B : \phi x \in U\}$, with $x \in B$ and $U \subset D$ open. Thus $i(B^*) \subset D^B$ has a subbase with sets $\{\phi \in i(B^*) : \phi x \in U\}$. On the other hand, B^* is a (topological) subspace of V^* with the weak* topology generated by the seminorms $\|\cdot\|_x$ for $x \in V$, and so subbasic open sets are of the form

$$B_x(\phi,\epsilon) \cap B^* = \{\psi \in B^* : |\psi x - \phi x| < \epsilon\} = \{\psi \in B^* : \psi x \in B(\phi x, \epsilon)\}.$$

Suppose that $\psi \in B^*$ is such that $\psi \in i^{-1}\{\phi \in i(B^*) : \phi x \in U\}$; thus, $\psi x = \psi|_B x \in U$. Since $U \subset D$ is open[10], there exists $\epsilon > 0$ such that

$$\psi x \in B(\psi x, \epsilon) \cap D \subset U.$$

Thus,

$$\psi \in B_x(\psi, \epsilon) \cap B^* \subset i^{-1}\{\phi \in i(B^*) : \phi x \in U\}$$

(check this!), which proves the continuity of i.

Now suppose $\psi \in i(B_x(\phi, \epsilon) \cap B^*)$; that is, ψ is the restriction to B of a map $\tilde{\psi} \in B^*$ satisfying $\psi x = \tilde{\psi} x \in B(\phi x, \epsilon)$; thus, there exists $\epsilon' > 0$ such that $B(\psi x, \epsilon') \subset B(\phi x, \epsilon)$. It follows that (check this!)

$$\psi \in \{\psi' \in i(B^*) : \psi' x \in B(\psi x, \epsilon')\} \subset i(B_x(\phi, \epsilon) \cap B^*),$$

which implies that i^{-1} is continuous. Therefore i is a homeomorphism and we are done.

[10] This means that it is the intersection of an open subset of $\mathbf{C}$ with D.

Also, one easily shows that B^* is closed in D^B. [See Brezis, Theorem 3.16.] The idea is to write B^* as a big intersection of compact and/or closed subsets of D^B that encode the conditions under which an arbitrary map $\phi \in D^B$ happens to be an element of B^*. This intersection consists of three types of sets.

(1) Sets that encode the boundedness of operator norm. Here we have $\{\phi \in D^B : |\phi(x)| \le \|x\|_V \text{ for all } x \in B\}$, which guarantees that $\|\phi\|_{V^*} \le 1$. This set can be written as a product $\prod_{x \in B} B(0, \|x\|)$ of compact balls, which is itself compact by Tychonoff's theorem.

(2) Sets that encode additivity. We have sets of the form $\{\phi \in D^B : \phi(x+y) - \phi(x) - \phi(y) = 0\}$ for all $x, y \in B$. To show that sets are closed, we show that $\phi \mapsto \phi(x+y) - \phi(x) - \phi(y)$ is continuous. We know that $\phi(x+y)$ is the projection $\pi_{x+y}(\phi) \colon D^B \to D$, which is continuous by definition of the product topology, and similarly for $\phi(x)$ and $\phi(y)$. Since sums of continuous functions are continuous, this gives the claim. Let me elaborate somewhat. We can think of the mapping $\phi \mapsto f(\phi) + g(\phi)$ as the composition

$$\begin{array}{ccccccc} D^B & \xrightarrow{\Delta} & D^B \times D^B & \xrightarrow{f \times g} & \mathbf{C} \times \mathbf{C} & \xrightarrow{+} & \mathbf{C} \\ \phi & \longmapsto & (\phi, \phi) & \longmapsto & (f(\phi), g(\phi)) & \longmapsto & f(\phi) + g(\phi). \end{array}$$

The diagonal map Δ is continuous for arbitrary topological spaces. The product map $f \times g$ of continuous maps is always continuous. Thus, given continuous functions $f, g \colon X \to \mathbf{C}$ out of an arbitrary topological space X, we see that $f + g$ is continuous.

(3) Sets that encode the scalar multiplication property. We have sets of the form $\{\phi \in D^B : \phi(\lambda x) - \lambda\phi(x) = 0\}$ for all $\lambda \in \mathbf{C}$ and $x \in B$.

Exercise 1.9.24.

Exercise 1.9.25.

Exercise 1.9.26. It suffices to prove that the weak operator topology is Hausdorff. Recall that the weak operator topology has a subbase generated by balls of the form

$$B_{x,\lambda}(T,\epsilon) := \{T' \in B(X \to Y) : |\lambda(T'x - Tx)| < \epsilon\}.$$

Let $T \in B(X \to Y)$ be non-zero; we will find balls that separate 0 and T. Choose $x \in X$ such that $Tx = 1$, and choose $\lambda \in Y^*$ such that $\lambda Tx = 1$ (using the Hahn–Banach theorem). Then $B_{x,\lambda}(0, 1/2)$ and $B_{x,\lambda}(T, 1/2)$ are disjoint, since we would have $|\lambda Tx| < 1$ otherwise by the triangle inequality.

Exercise 1.9.27. (i) Suppose that $\|T_n\|_{\mathrm{op}} \to 0$. By the Cauchy–Schwarz inequality, we have

$$|\langle T_n x_n, y_n \rangle| \leq \|T_n x_n\|_H \|y_n\|_H \leq \|T_n\|_{\mathrm{op}} \|x_n\|_H \|y_n\|_H,$$

which goes to zero as x_n and y_n are bounded sequences.

Conversely, suppose $\|T_n\|_{\mathrm{op}} \not\to 0$. Then, there exists $\epsilon > 0$ such that $\|T_{n_j}\|_{\mathrm{op}} > \epsilon$ along some subsequence; thus there exists a sequence x_j with $\|x_j\|_H = 1$ and $\|T_{n_j} x_j\|_H > \epsilon$. It follows that

$$\left\langle T_{n_j} x_j, \frac{T_{n_j} x_j}{\|T_{n_j} x_j\|_H} \right\rangle = \|T_{n_j} x_j\|_H \not\to 0$$

as needed.

(ii)

(iii)

(iv)

Exercise 1.9.28.

Exercise 1.9.29.

Exercise 1.9.30.

1.10. Continuous functions on locally compact Hausdorff spaces

Lovers and madmen have such seething brains,
Such shaping fantasies, that apprehend
More than cool reason ever comprehends.
The lunatic, the lover, and the poet
Are of imagination all compact.

— Theseus, in *A Midsummer Night's Dream* Act 5 Scene 1 (c. 1600)

Exercise 1.10.1. It suffices to verify that the map

$$x \mapsto \frac{d(x,K)}{d(x,K)+d(x,L)}$$

is continuous; in turn, it suffices to show that $x \mapsto d(x,K)$ is Lipschitz continuous. If $d(x,K) > d(x,y) + d(y,K)$ for some x,y, then $d(x,k) \geq d(x,K) > d(x,y) + d(y,k)$ for some $k \in K$, violating the triangle inequality. Thus $d(x,K) - d(y,K) \leq d(x,y)$ for all x,y as needed.

Exercise 1.10.2. (i) Suppose $K \subset X$ is compact and $x \notin K$. Then, for each $y \in K$, there exist disjoint open sets $y \in U_y$ and $x \in V_y$. By compactness, we obtain a finite subcover $K \subset U_{y_1} \cup \cdots \cup U_{y_n}$. It follows that $x \in V_{y_1} \cap \cdots \cap V_{y_n}$, and thus we obtain open sets separating K and x as needed.

(ii) Suppose $K, L \subset X$ are compact. Using (i), for each $x \in L$ we obtain disjoint open sets $K \subset U_x$ and $x \subset V_x$. By compactness of L we obtain a finite subcover $L \subset V_{x_1} \cup \cdots \cup V_{x_n}$; since $K \subset U_{x_1} \cap \cdots \cap U_{x_n}$ the result follows.

(iii) This follows from the fact that closed subsets of compact spaces are compact.

Exercise 1.10.3. The space $(\mathbf{R}, \mathcal{F}')$ is Hausdorff since it is stronger than the usual Hausdorff topology on $\mathbf{R}$. Similarly every point is closed. This space is not normal however — notice that every open set is of the form W, $W \cup \mathbf{Q}$, or $W \cap \mathbf{Q}$ for $W \in \mathcal{F}$, and consider the disjoint closed sets $\mathbf{R} \setminus \mathbf{Q}$ and $\{0\}$. Suppose $\mathbf{R} \setminus \mathbf{Q} \subset U$ and $0 \in V$ are open sets in $\mathcal{F}'$; let us consider cases on the form of V. If $V \in \mathcal{F}$, then $0 \in (-\epsilon, \epsilon) \subset V$ for some $\epsilon > 0$; thus $V \cap (\mathbf{R} \setminus \mathbf{Q})$ is non-empty. If $V = V' \cup \mathbf{Q}$ for some $V' \in \mathcal{F}$, then this can be reduced to the first case. Finally, if $V = V' \cap \mathbf{Q}$ for some $V' \in \mathcal{F}$, then $0 \in (-\epsilon, \epsilon) \cap \mathbf{Q} \subset V$ for some $\epsilon > 0$. Choose $x \in (-\epsilon, \epsilon) \setminus \mathbf{Q} \subset \mathbf{R} \setminus \mathbf{Q}$; we must have $x \in U' \subset U$ for some $U' \in \mathcal{F}'$. We see that case 1 and 2 require U' to contain some elements of $(-\epsilon, \epsilon) \cap \mathbf{Q}$, and case 3 is not possible. Thus we see that in all cases, $\mathbf{R} \setminus \mathbf{Q}$ and $\{0\}$ cannot be separated by open sets in $\mathcal{F}'$.

Exercise 1.10.4. (i) The set $\mathbf{N}$ is discrete and thus Hausdorff. The product of Hausdorff spaces is Hausdorff (exercise 1.8.18); thus $\mathbf{N}^{\mathbf{R}}$ is Hausdorff. Since $\{(n_x)_{x \in \mathbf{R}}\} = \bigcap_{x \in \mathbf{R}} \pi_x^{-1}(\{n_x\})$, we see that points are closed (so $\mathbf{N}^{\mathbf{R}}$ is T_1).

(ii) Given $(n_x)_{x \in \mathbf{R}} \in \mathbf{N}^{\mathbf{R}}$, it is not possible for all but countably many of its components to be equal to 1 and all but countably many of its components to be equal to 2 at the same time, since $\mathbf{R}$ is uncountable. Thus K_1 and K_2 are disjoint. Now suppose $(n_x)_{x\in\mathbf{R}}$ is adherent to K_1. It suffices to prove that we cannot have $n_y = n_{y'} = k \geq 2$ for $y \neq y'$. Indeed, if this were the case, then there would exist a point of K_1 belonging to $\pi_y^{-1}(\{k\}) \cap \pi_{y'}^{-1}(\{k\})$, contradicting injectivity for K_1.

(iii)

(iv)

(v)

Exercise 1.10.5.

Exercise 1.10.6. Since X is locally compact, each point $x \in K$ admits an open neighborhood V_x with compact closure $\overline{V_x}$; by compactness of K, there exists a finite subcover $V := V_{x_1} \cup \cdots \cup V_{x_n}$ of K. Notice that $\overline{V}$ is then a compact neighborhood of K. Let $U' := U \cap V \supset K$; then $\overline{U'}$ is compact, since it is a closed subset of the intersection of a closed and compact set. It follows that $\overline{U'}$ is a compact Hausdorff neighborhood of K, and is thus normal. We conclude that there exists a continuous compactly supported function $f \colon \overline{U'} \to \mathbf{R}$ such that $1_K \leq f \leq 1_{U'}$; we may extend f to X by defining it to be 0 outside U'.

This is somewhat tangential, but here's a cool argument using nets proving that the intersection of a closed set L and a compact set K is itself compact.

We show that every net in $L \cap K$ has a convergent subnet. Let $(x_\alpha)_{\alpha\in A}$ be a net in $L \cap K$. Then, by compactness of K, there exists a convergent subnet $(x_{\phi(\beta)})_{\beta\in B}$ converging to a limit $x \in K$. Therefore, x is an adherent point of L, and so $x \in K \cap L$ as needed.

Exercise 1.10.7. For now, let us assume that every non-empty open set has positive measure. We first prove that

$$\overline{C_c(X \to \mathbf{R})} \subset C_0(X \to \mathbf{R}).$$

Suppose $f \in L^\infty(X,\mu) \setminus C_0(X \to \mathbf{R})$. Then there exists $\epsilon > 0$ such that, for every compact $K \subset X$, we have $|f(x)| > \epsilon$ for some $x \in X \setminus K$. Given $g \in C_c(X \to \mathbf{R})$ supported on some compact K, we have $\|f - g\|_{L^\infty} > \epsilon$, since $\{y \in X \setminus K : |f(y)| > \epsilon\}$ is a non-empty open set and thus has positive measure by hypothesis.

Now let $f \in C_0(X \to \mathbf{R})$.

For the general case, we will work on the *support* of μ, which is defined by

$$\operatorname{supp}(\mu) := \{x \in X : \text{every neighborhood } U \text{ of } x \text{ satisfies } \mu(U) > 0\}.$$

Notice that $\operatorname{supp}(\mu)$ is closed, since its complement is the union of open sets (of measure zero). Also notice that every non-empty open subset of $\operatorname{supp}(\mu)$ is of positive measure. ...

Exercise 1.10.8.

Exercise 1.10.9.

Exercise 1.10.10. [Relied on `https://math.stackexchange.com/a/1188995/` for some details I missed.] Notice that f is bounded, and K is closed. As in exercise 1.10.6, there exists an open neighborhood U of K with compact closure; use the Tietze extension theorem to extend f to $\overline{U}$. If U is clopen in X, then $1_U f$ works. Otherwise we apply

Urysohn's lemma to obtain $g\colon \overline{U} \to [0,1]$ satisfying $1_K \leq g \leq 1_U$; then $gf1_{\overline{U}}$ works.

Exercise 1.10.11. If 1_E is lower semicontinuous, then the set $E = f^{-1}((1/2, +\infty))$ is open. Conversely, $f^{-1}((a, +\infty))$ equal to either $\varnothing$, E, or X, depending on a. The result for upper semicontinuity is proven analogously.

Suppose X is normal Hausdorff and f is upper semicontinuous. Then ...