245B Real Analysis: Some solutions

ho boon suan

January 2, 2022 to February 1, 2022, 03:27

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

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

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 $\int_{x\in E:|f(x)|>\lambda} |f|\,d\mu \le \epsilon/2$. It follows that

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

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

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 $E_1, E_2, \ldots$, 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 $\{a_1, \ldots, a_n\}$. We write $g = \sum_{i=1}^n a_i 1_{g^{-1}(\{a_i\})}$, and compute

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

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 $-\nu \le \mu \le \nu$, or

$$-\nu_+ + \nu_- \le \mu_+ - \mu_- \le \nu_+ - \nu_-.$$

Our goal is to prove that $|\mu| \le \nu$, or $\mu_+ + \mu_- \le \nu_+ - \nu_-$. 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 $\mu = \mu_{\mathrm{ac}} + \mu_{\mathrm{s}}$ uniquely, with $\mu_{\mathrm{ac}} \ll m$ and $\mu_{\mathrm{s}} \perp m$. We will further decompose

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

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

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

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

(1) E is at most countable.

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

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

We first prove (1). Suppose for contradiction that E is uncountable. Since $\mu_{\mathrm{s}} \le \mu$ 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 $\mu_{\mathrm{s}}(X_n) < \infty$, with X_n disjoint. Then, $E = \bigcup_{n=1}^\infty (E \cap X_n)$, with $\mu_{\mathrm{s}}(E \cap X_n) < \infty$ 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

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

for $k \ge 2$, with $E_{n,1} := \{x \in E \cap X_n : \mu_s(\{x\}) \ge 1\}$. 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

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

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 $\mu_{\mathrm{s}} \perp m$. Indeed, $\mu_{\mathrm{sc}} \le \mu_{\mathrm{s}}$, 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 $\sum_{i=1}^\infty |f(y_i) - f(x_i)| \le \epsilon$ 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 $\sum_{i=1}^n |f(y_i) - f(x_i)| \le \epsilon$

for all n. Thus we conclude that $\sum_{i=1}^{\infty} |f(y_i) - f(x_i)| \leq \epsilon$ 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,

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

for all $n \geq 1$, and so

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

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

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

as needed.

Conversely, suppose $x \mapsto \mu([0, x])$ is continuous. Fix $x \in [0, +\infty]$. We prove that $\mu(\{x\}) \leq \epsilon$ 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

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

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 $m(E) \leq \delta$. Suppose $[x_1, y_1], \ldots, [x_n, y_n]$ are disjoint intervals in $[0, +\infty]$ of total length at most δ, so that $m(\bigcup_{i=1}^n (x_i, y_i]) \leq \delta$. Then

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

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 $[x_1, y_1], \ldots$ 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 $\mu(E) \leq \mu(U) < \epsilon$. 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

$$d(f,h) = \|f-h\| \leq \|f-g\| + \|g-h\| = d(f,g) + d(g,h).$$

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

$$\begin{aligned}\|f+g\| &= d(0,f+g)\\ &\leq d(0,f) + d(f,f+g)\\ &= d(0,f) + d(0,g)\\ &= \|f\| + \|g\|.\end{aligned}$$

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 $m, n \geq N$, we have by repeated applications of the triangle inequality

$$\begin{aligned}\Big\|\sum_{j=m}^{n} f_j\Big\| &\leq \sum_{j=m}^{n} \|f_j\|\\ &\leq \sum_{j=N}^{\infty} \|f_j\|\\ &< \epsilon.\end{aligned}$$

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 $N_1 < N_2 < N_3 < \ldots$ such that $\|f_m - f_n\| \leq \epsilon/2^k$ whenever $m, n \geq N_k$. Then $\sum_{j=k}^{\infty} \|f_{N_j} - f_{N_{j-1}}\| \leq \epsilon/2^{k-1}$, where $k \geq 2$. 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 $1 \leq p < \infty$. 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

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

as $n \to \infty$.

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

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

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

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

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

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

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.

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

Suppose now that the vectors x_i are linearly dependent. Then we may write $z_1x_1 + \cdots + z_nx_n = 0$ for some $z = (z_1, \ldots, z_n) \in \mathbf{C}^n \setminus \{0\}$, 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 $e_1, \ldots, e_n$. 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 $\{e_1, \ldots, e_{n+1}\}$ is orthonormal, with span equal to $\{e_1, \ldots, e_n, v\}$.

Therefore, given some n-dimensional complex inner product space V, we may take a basis $\{e'_1, \ldots, e'_n\}$ 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

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

this quantity can be made arbitrarily small for $m, n \ge N$ 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

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

whenever $m, l \ge N$. 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 $\|f_{2,m} - f_{2,l}\| \le 1/2$ whenever $m, l \ge n_2$. 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 $\|f_{k,m} - f_{k,l}\| \le 1/k$ whenever $m, l \ge n_k$ and $f_k := f_{k,n_k}$. The sequence $(f_n)_{n=1}^{\infty}$ is Cauchy since, for sufficiently large l and m with $l \ge m$, we have by $(*)$:

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

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

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 $\|f_n - x\| \le \epsilon/2$. This f_n is in turn a countable sum $\sum_{k=1}^\infty c_{n,k} e_{n,k}$, and so we have

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

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;*
悟無好惡 *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 $\{e_1, \dots, e_n\}$ 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

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

(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 $f(c_1, \dots, c_n) := \|\sum_{i=1}^n c_i e_i\|_X$. Then f is continuous, as we may compute (with $C := \max_{1\le i\le n} \|e_i\|_X$)

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

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 $\|\cdot\|_{\mathrm{op}} := \|\cdot\|_{B(X\to Y)}$ is a norm. If $\|T\|_{\mathrm{op}} = 0$, 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 $\|0\|_{\mathrm{op}} = 0$.

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

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

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

Finally, we have

$$\|Sx + Tx\|_Y \le \|Sx\|_Y + \|Tx\|_Y \le \|S\|_{\text{op}}\|x\|_X + \|T\|_{\text{op}}\|x\|_X$$

for all $x \in X$, so that $\|S+T\|_{\text{op}} \le \|S\|_{\text{op}} + \|T\|_{\text{op}}$ 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 $\|T_m - T_n\|_{\text{op}} \le \epsilon$ whenever $m, n \ge N$. That is, $\|T_m x - T_n x\|_Y \le \epsilon\|x\|_X$ whenever $m, n \ge N$ 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 $\lim_{n\to\infty}\|T_n - T\|_{\text{op}} = 0$.

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

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

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

Finally, we prove that $\lim_{n\to\infty}\|T_n - T\|_{\text{op}} = 0$. It suffices to prove that $\|T_n x - Tx\|_Y \le \epsilon\|x\|_X$ 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 $\|T_n x - Tx\|_Y \le \epsilon$ for large n. Finally, if $\|x\|_X \ne 0$, the result follows from the $\|x\|_X = 1$ case together with homogeneity.

Exercise 1.5.3. We compute

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

which implies that $\|ST\|_{\text{op}} \le \|S\|_{\text{op}}\|T\|_{\text{op}}$.

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 $\|f\|_{\text{op}} = \|\overline{f}\|_{\text{op}}$.

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. □

(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 $|\int_X fg\, d\mu| \leq \|f\|_{L^p}\|g\|_{L^{p'}}$, and so $\|Tg\|_{(L^p)^*} \leq \|g\|_{L^{p'}}$. 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

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

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

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

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

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:

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

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, suppose $x \in X$ is such that $\lambda(x) = 0$ for all $\lambda \in Y^\perp$. Then $\iota(x)(\lambda) = 0$ for all $\lambda \in Y^\perp$, so that $\iota(x) \in (Y^\perp)^\perp$. ...

(ii)

(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.

Exercise 2.4.2.

Exercise 2.4.3.

Exercise 2.4.4.

Exercise 2.4.5.

Exercise 2.4.6.

Exercise 2.4.7.

Exercise 2.4.8.

Exercise 2.4.9.

Exercise 2.4.10.

Exercise 2.4.11.

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.

Exercise 1.6.2.

Exercise 1.6.3.

Exercise 1.6.4.

Exercise 1.6.5.

Exercise 1.6.6.

Exercise 1.6.7.

Exercise 1.6.8.

Exercise 1.6.9.

Exercise 1.6.10.

Exercise 1.6.11.

Exercise 1.6.12.

Exercise 1.6.13.

Exercise 1.6.14.

Exercise 1.6.15.

Exercise 1.6.16.

Exercise 1.6.17.