TERENCE TAO'S "AN EPSILON OF ROOM" CHAPTER 4 EXERCISES

KELLER VANDEBOGERT

1. Exercise 1.4.1

Assume first that $\langle Tx, Ty\rangle = \langle x, y\rangle$. Then,

$$||Tx||^2 = ||x||^2 \implies ||Tx|| = ||x||$$

So T is an isometry.

Conversely, suppose that T is an isometry. In the complex case, employ the polarization identities to see

$$\begin{aligned}\langle Tx, Ty\rangle &= \frac{1}{4}\Big(|\ Tx+Ty||^2 - ||Tx-Ty||^2 + i||Tx+iTy||^2 - i||Tx-iTy||^2\Big)\\ &= \frac{1}{4}\Big(||x+y||^2 - ||x-y||^2 + i||x+iy||^2 - i||x-iy||^2\Big)\\ &= \langle x, y\rangle\end{aligned}$$

So that $\langle Tx, Ty\rangle = \langle x, y\rangle$. In the real case, the polarization identities are even simpler:

$$\begin{aligned}\langle Tx, Ty\rangle &= \frac{1}{4}\Big(|\ Tx+Ty||^2 - ||Tx-Ty||^2\Big)\\ &= \frac{1}{4}\Big(||x+y||^2 - ||x-y||^2\Big)\\ &= \langle x, y\rangle\end{aligned}$$

Which gives the result.

Date: September 17, 2017.

2. Exercise 1.4.2

Note that

$$\left(\overline{\langle x_i, x_j\rangle}\right)^t = \left(\overline{\langle x_j, x_i\rangle}\right)$$
$$= \left(\langle x_i, x_j\rangle\right)$$

So this matrix is Hermitian. Let $X := (x_1\ x_2\ \dots\ x_n)$, the x_i are our column vectors. It is trivial that

$$\overline{X}^t X = \left(\langle x_i, x_j\rangle\right)$$

which is our Gram matrix. For any nonzero v, $Xv \neq 0$ whenever the x_i are linearly independent. Thus,

$$\overline{v}^t\overline{X}^t Xv = \overline{Xv}^t Xv$$
$$= \langle Xv, Xv\rangle$$
$$= ||Xv||^2 > 0$$

If we suppose that x_i are linearly dependent, then X has nontrivial kernel. Choose $v \in \operatorname{Ker} X$, so that

$$\overline{v}^t\overline{X}^t Xv = \overline{Xv}^t Xv = 0$$

So the matrix is positive semidefinite.

3. Exercise 1.4.3

By induction, with base case $n = 2$:

$$||x_1 + x_2||^2 = ||x_1||^2 + 2\mathrm{Re}\langle x_1, x_2\rangle + ||x_2||^2$$
$$= ||x_1||^2 + ||x_2||^2$$

Now assume $n > 2$, where the inductive hypothesis holds for all integers less than n. It is obvious that if our vectors x_i are orthogonal, in particular, x_n and $x_1 + \cdots + x_{n-1}$ are orthogonal. Using this:

$$\begin{aligned} ||x_1 + x_2 + \cdots + x_{n-1} + x_n||^2 &= ||x_1 + \cdots + x_{n-1}||^2 + 2\text{Re}\langle x_1 + \cdots + x_{n-1}, x_n\rangle + ||x_n||^2 \\ &= ||x_1 + \cdots + x_{n-1}||^2 + ||x_n||^2 \\ &= ||x_1||^2 + ||x_2||^2 + \cdots + ||x_n||^2 \end{aligned}$$

Whence the result.

4. Exercise 1.4.4

Let $\{e_\alpha\}_{\alpha \in A}$ denote our orthonormal basis. Suppose we have a linear combination with constants c_α such that

$$c_1 x_1 + \ldots c_n x_n = 0$$

Taking the inner product in the above with each e_i, we immediately deduce that $c_i = 0$ for all i, so this set is linearly independent. Given any x, there exist constants c_i such that

$$x = \sum_{i=1}^{n} c_i e_i$$

Again, taking the inner product in the above with each of the e_i, we find that

$$\langle x, e_i \rangle = c_i$$

Implying

$$x = \sum_{i=1}^{n} \langle x, e_i \rangle e_i$$

as asserted.

5. Problem 1.4.5

Assume $(x_i)_{i=1}^n$ is orthonormal. Set

$$x_{n+1} := v - \sum_{i=1}^{n} \langle v, x_i \rangle x_i$$

Clearly $x_{n+1} \in \text{Span}\{x_1, \dots, x_n, v\}$, so that

$$\text{Span}\{x_1, \dots, x_n, v\} = \text{Span}\{x_1, \dots, x_n, x_{n+1}\}$$

And we also see for each $j \leqslant n$,

$$\begin{aligned} \langle x_{n+1}, x_j \rangle &= \langle v, x_j \rangle - \sum_{i=1}^{n} \langle v, x_i \rangle \langle x_i, x_j \rangle \\ &= \langle v, x_j \rangle - \langle v, x_j \rangle = 0 \end{aligned}$$

So we have produced a larger orthonormal set (upon normalizing x_{n+1} as needed). Hence given any basis $\{b_1, \dots, b_n\}$, we can produce an orthonormal basis $\{x_1, \dots, x_n\}$ by setting $b_1 = x_1$, and then inductively defining our x_i by the above process. By construction,

$$\text{Span}\{b_1, \dots, b_n\} = \text{Span}\{x_1, \dots, x_n\}$$

So this is indeed a basis.

6. Exercise 1.4.6

For the parallelogram law,

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

Suppose now that $p \neq 2$. Choose disjoint measurable sets E_1, E_2. Then,

$$\begin{aligned} ||\chi_{E_1} + \chi_{E_2}||_p^2 + ||\chi_{E_1} - \chi_{E_2}||_p^2 &= 2\big(\mu(E_1) + \mu(E_2)\big)^{2/p} \\ &\neq 2\mu(E_1)^{2/p} + 2\mu(E_2)^{2/p} \\ &= 2||\chi_{E_1}||_p^2 + ||\chi_{E_2}||_p^{2/p} \end{aligned}$$

So the parallelogram law does not hold for $p \neq 2$. Now, in order to prove the Hanner inequalities, we may assume without loss of generality that $||f||_p = 1$ and $||g||_p < 1$ by homogeneity. Using the inequality supplied by the hint, we may assume $f,\ g \geqslant 0$ by splitting into positive/negative parts of course. We see:

$$|f+g|^p + |f-g|^p \geqslant \Big((1+||g||_p)^{p-1} + (1-||g||_p)^{p-1}\Big) \cdot f^p + \Big((1+||g||_p)^{p-1} - (1-||g||_p)^{p-1}\Big) \cdot ||g||_p^{1-p} g^p$$

Integrating the above yields

$$\begin{aligned} ||f+g||_p^p + ||f-g||_p^p &\geqslant \Big((1+||g||_p)^{p-1} + (1-||g||_p)^{p-1}\Big) \\ &\quad + \Big((1+||g||_p)^{p-1} - (1-||g||_p)^{p-1}\Big) \cdot ||g||_p \\ &= (1+||g||_p)^p + (1-||g||_p)^p \end{aligned}$$

Which establishes the first inequality. For the second inequality, we want to employ the above. Putting $f \mapsto f+g$, $g \mapsto f-g$, the first inequality yields

$$\begin{aligned} \Big(||f+g||_p + ||f-g||_p\Big)^p + \big|||f+g||_p - ||f-g||_p\big|^p &\leqslant ||2f||_p^p + ||-2g||_p^2 \\ &= 2^p\Big(||f||_p^p + ||g||_p^2\Big) \end{aligned}$$

Which yields the result.

7. Exercise 1.4.7

A subspace of a Hilbert space is a Hilbert space if and only if it contains all of its limit points, since the rest of the structure is inherited by the ambient space. But a subspace is closed if and only if it contains all of its limit points, so the result follows trivially.

8. Exercise 1.4.8

We define the completion by the union of V with all limits of Cauchy sequences in V. By continuity of inner products, it is well defined to set

$$\langle x, y\rangle := \lim_{n\to\infty} \langle x_n, y_n\rangle, \quad x_n \to x,\ y_n \to y$$

This space is trivially complete by construction, so we are done.

9. Exercise 1.4.9

The vector space structure and completeness are both obvious. It remains only to show the inner product properties. One sees:

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

Conjugate linearity and linearity are trivial. Also,

$$\begin{aligned}\langle (x,x'),(x,x')\rangle_{H\oplus H'} &= \langle x,x\rangle_H + \langle x',x'\rangle_{H'} \\ &= ||x||_H^2 + ||x'||_{H'}^2 \\ &= 0 \iff x,\ x' = 0 \\ &\iff (x,x') = 0\end{aligned}$$

Which proves nondegeneracy.

10. Exercise 1.4.10

Consider $B := \{x : ||x|| < 1\}$ in $\mathbb{R}^2$. Then, no point outside of B has a minimizer, and B is open. Now consider

$$B_1 := \{x \mid ||x|| \leqslant 1\}, \quad B_2 := \{x \mid ||x-(4,0)|| \leqslant 1\}$$

Then $B_1 \cup B_2$ is closed but not convex. Consider then the point $(2,0)$. The minimum distance from $B_1 \cup B_2$ is 1, achieved by the points $(1,0)$ and $(3,0)$, so these minimizers are not unique.

For the case of a pre-Hilbert space, consider $C([0,1]) \subset L^2([0,1])$. Set K to be the continuous functions supported on $[0,1/2]$, and consider the distance of $\chi_{[0,1]}$ from K. It is easy to see that $d(\chi_{[0,1]},K) = \frac{1}{\sqrt{2}}$, and is achieved by the sequence of functions

$$\eta_{1/n} * \chi_{[0,1/2]}$$

As $n \to \infty$, where η_ϵ denotes the standard mollifier and $*$ the operation of convolution. Noting that $[0, 1/2]$ is a compact set, $\eta_{1/n} * \chi_{[0,1/2]} \to \chi_{[0,1/2]}$ uniformly, but this limit is not contained in $C([0, 1])$.

Suppose now that K is compact, and define $D := \inf_{y \in K} ||x - y||$. Choose a sequence y_n such that $\lim_{n\to\infty} ||x-y_n|| = D$. By compactness, y_n has a convergent subsequence $y_{n_k} \to y$. Since K is closed, $y \in K$, so the minimizer exists.

11. Exercise 1.4.11

The case $p = 2$ is already solved. Suppose first that $p > 2$. Choosing a sequence y_n such that $||y_n - x|| \to D$, Hanner's inequality yields ($y_n \in K$, K closed/compact):

$$\begin{aligned} ||2\Big(x - \frac{y_n + y_m}{2}\Big)||_p^p + ||y_n - y_m||_p^p \leqslant \Big(||y_n - x||_p + ||y_m - x||_p\Big)^p \\ + \Big(||y_n - x||_p - ||y_m - x||_p\Big)^p \end{aligned}$$

Letting $m,\ n \to \infty$, we see:

$$\begin{aligned} \lim_{m,\ n\to\infty} ||y_n - y_m||_p^p &\leqslant 2^p D + (D - D)^p - 2^p D^p \\ &= 0 \end{aligned}$$

So that our sequence is Cauchy, hence convergent to some limit $y_n \to y \in K$.

Now consider the case $p \in (1, 2)$. Hanner's inequalities are reversed in this case, and without loss of generality we may assume that $x = 0$. Again, choose y_n such that $||y_n||_p \to D$, our minimizing distance. We see:

$$\Big(||y_n+y_m||_p+||y_n-y_m||_p\Big)^p+\Big(||y_n+y_m||_p-||y_n-y_m||_p\Big)^p$$
$$\leqslant 2^p\Big(||y_n||_p^p+||y_m||_p^p\Big)$$

By convexity of K, $||y_n+y_m||_p \to 2D$. Suppose for sake of contradiction that there exists from $\epsilon > 0$ such that $||y_n - y_m||_p \to \epsilon$. Letting $m,\ n \to \infty$, the above inequality gives

$$|2D+\epsilon|^p+|2D-\epsilon|^p \leqslant 2^pD^p+2^pD^p$$

However, if $f(x) := |2D+x|^p$, f is strictly convex, since $p > 1$. But the above says that

$$\frac{1}{2}f(\epsilon)+\frac{1}{2}f(-\epsilon) \leqslant f(0)$$

whereas the reverse inequality holds by convexity of f. Hence, we deduce

$$\frac{1}{2}f(\epsilon)+\frac{1}{2}f(-\epsilon) = f(0)$$

By strict convexity, this is possible if and only if $\epsilon = 0$. Hence, y_n is Cauchy and must converge to some minimizer $y_n \to y \in K$. This completes the proof.

12. Exercise 1.4.12

Let $\{x_1,\dots,x_n\}$ be an orthonormal basis of V. Define $x_V := \sum_{i=1}^n \langle x, x_i\rangle x_i$. Setting $x_{V^\perp} := x - x_V$, we see that

$$\langle x_{V^\perp}, x_i\rangle = \langle x, x_i\rangle - \langle x_V, x_i\rangle$$
$$= \langle x, x_i\rangle - \langle x, x_i\rangle = 0$$

So that $x_{V^\perp}$ is orthogonal to every element of V. To see that x_V minimizes distances, notice that for any $y \in V$:

$$x-y = x-x_V+x_V-y$$
$$\implies ||x-y||^2 = ||x-x_V||^2+||y-x_V||^2 \geqslant ||x-x_V||^2$$

So that x_V is indeed our minimizer.

13. Exercise 1.4.13

(a). Note first that $\langle \cdot, v \rangle$ is a continuous functional. This follows by the Cauchy=Schwarz inequality. Hence,

$$V^\perp = \bigcap_{v \in V} f_v^{-1}(\{0\})$$

where $f_v(x) := \langle v, x \rangle$. By continuity, $f_v^{-1}(\{0\})$ is closed, and $V^\perp$ is the intersection of closed sets, hence closed.

We proceed to show $\overline{V} = \left(V^\perp\right)^\perp$. Since $V \subset \left(V^\perp\right)^\perp$, we trivially have that $\overline{V} \subset \left(V^\perp\right)^\perp$.

For the reverse inclusion, suppose for sake of contradiction that the above containment is proper. Since $\overline{V}$ is closed, $\overline{V}^\perp$ is nontrivial. Choose $v \in \overline{V}^\perp$, and observe that $\overline{V}^\perp \subset V^\perp$ (since $V \subset \overline{V}$).

But this implies that $v \in V^\perp$, and $v \in \left(V^\perp\right)^\perp$, so

$$v \in V^\perp \cap \left(V^\perp\right)^\perp = \{0\}$$
$$\implies v = 0$$

Contradicting our assumption. We conclude that $\overline{V} = \left(V^\perp\right)^\perp$.

(b). If $V^\perp = \{0\}$, then by part (a), $\overline{V} = \left(V^\perp\right)^\perp = H$, so that V is dense.

Conversely, if V is dense, $\overline{V} = H \implies \overline{V}^\perp = \{0\}$. But it is obvious that $\overline{V}^\perp = \overline{V^\perp} = V^\perp$ since $V^\perp$ is closed. Therefore $V^\perp = \{0\}$.

(c). By Exercise 1.4.12, we can decompose $x = x_V + x_{V^\perp}$. Then, clearly $H = V + V^\perp$. As $V \cap V^\perp = \{0\}$, we conclude that $H = V \oplus V^\perp$.

(d). Let $x \in (V+W)^\perp$. Then, $\langle v+w, x\rangle = 0$ for all $v \in V$, $w \in W$. Setting $v = 0$, we see $\langle w, x\rangle = 0$ for all $w \in W$, and similarly, $\langle v, x\rangle = 0$ for all $v \in V$, so that $x \in V^\perp \cap W^\perp$.

Conversely, if $x \in V^\perp \cap W^\perp$, then for all $v \in V$, $w \in W$,

$$\langle v+w, x\rangle = \langle v, x\rangle + \langle w, x\rangle = 0$$

So $x \in (V+W)^\perp$.

Now, recalling that $\overline{V}^\perp = V^\perp$, and likewise for W, we see

$$\begin{aligned}&(V^\perp + W^\perp)^\perp = \overline{V} \cap \overline{W} \\ &\implies \overline{(V^\perp + W^\perp)} = \overline{V}^\perp \cap \overline{W}^\perp = V^\perp \cap W^\perp\end{aligned}$$

Whence the result.

14. Exercise 1.4.14

Set $K = \mathrm{Ker}(\lambda)$. This is closed by continuity and trivially convex. If $\lambda \equiv 0$, the result is obvious, so suppose $\lambda \not\equiv 0$ and choose $f \notin K$. We can find $h \in K$ such that $||f-h||_p$ is minimized by Exercise 1.4.11. Set $u := |f-h|^{p-2}(\overline{f} - \overline{g})$. For any $k \in K$, we have that

$$\mathrm{Re}\Big(\int_X uk d\mu\Big) \leqslant 0$$

However, by linearity, $k \in K \implies -k,\ ik \in K$, so we can substitute those in the above inequality to find that $\int_X uk d\mu = 0$ precisely. We also see

$$\begin{aligned}||u||_{p'}^{p'} &= \int_X \big(|f-h|^{p-2}|\overline{f} - \overline{h}|\big)^{\frac{p}{p-1}} d\mu \\ &= \int_X |f-h|^p d\mu \\ &= ||f-h||_p^p < \infty\end{aligned}$$

Thus $u \in L^{p'}$, since $f,\ h \in L^p$. Let $g \in L^p$. Now, decompose $g = g_1 + g_2$ where

$$g_1 := \frac{\lambda(g)}{\lambda(f)}(f-h), \quad g_2 := g - g_1$$

Note that

$$\lambda(g_2) = \lambda(g) - \lambda(g) = 0$$

This implies that $g_2 \in K$, so that we see

$$\begin{aligned}\int_X gud\mu &= \int_X g_1 ud\mu \\ &= \frac{\lambda(g)}{\lambda(f)}\int_X (f-h)|f-h|^{p-2}|\overline{f}-\overline{h})d\mu \\ &= \frac{\lambda(g)}{\lambda(f)}\int_X |f-h|^p d\mu \\ &= \lambda(g)\frac{||f-h||_p^p}{\lambda(f)}\end{aligned}$$

Since $f \notin K$, we know $||f-h||_p \neq 0$, so we may define $\phi := \frac{u\cdot\lambda(f)}{||f-h||_p^p}$ Then $\phi \in L^{p'}$, and, by construction,

$$\lambda = \lambda_\phi$$

Which proves the result.

15. Exercise 1.4.15

Let $\{e_1, \dots\}$ denote our orthonormal basis. Define $\lambda(x) := \langle Tx, e_i\rangle$. We see that

$$|\lambda(x)| \leqslant ||T|| \cdot ||x||$$

So λ is bounded, hence continuous. By the Riesz Representation theorem, there exist $v_i \in H$ such that

$$\lambda(x) = \langle x, v_i\rangle$$

Define $T^*e_i := v_i$ and extend by linearity. This uniquely defines an operator $T^* : H' \to H$. Note that

$$\langle x, T^*y \rangle = \langle Tx, y \rangle \leqslant ||Tx|| \cdot ||y||$$

So that T^* is continuous. For linearity,

$$\begin{aligned} \langle x, T^*(cy+z) \rangle &= \langle Tx, cy+z \rangle \\ &= \overline{c}\langle x, T^*y \rangle + \langle x, T^*z \rangle \\ &= \langle x, cT^*y + T^*z \rangle \end{aligned}$$

Completing the proof.

16. Exercise 1.4.16

(a). Note that for $x \in H$, $y \in H'$,

$$\begin{aligned} \langle Tx, y \rangle &= \langle x, T^*y \rangle \\ &= \langle T^{**}x, y \rangle \end{aligned}$$

As x and y are arbitrary, we see that $T = T^{**}$.

(b). Recall that T is an isometry if and only if $\langle Tx, Ty \rangle = \langle x, y \rangle$. By definition of adjoint,

$$\langle Tx, Ty \rangle = \langle x, T^*Ty \rangle$$

and we conclude that T is an isometry if and only if $y = T^*Ty$ for all y, that is $T^*T \equiv \text{id}_H$.

(c). Suppose first that $T^*T = \text{id}_H$, $TT^* = \text{id}_{H'}$. Then T is a right invertible isometry, hence an isomorphism.

Conversely, let T be an isomorphism. Then T is an isometry, so by the previous part, $T^*T = \text{id}_H$. By surjectivity, for any $x' \in H'$, we can find $x \in H$ such that $Tx = x'$. Given $x',\ y' \in H'$:

$$\begin{aligned}\langle x', y' \rangle &= \langle Tx, y' \rangle \\ &= \langle x, T^* y' \rangle \\ &= \langle Tx, TT^* y' \rangle \\ &= \langle x', TT^* y' \rangle\end{aligned}$$

From which we conclude that $TT^* = \mathrm{id}_{H'}$, as asserted.

(d). One one hand, we see

$$\langle TSx, y \rangle = \langle x, (TS)^* y \rangle$$

On the other,

$$\begin{aligned}\langle TSx, y \rangle &= \langle Sx, T^* y \rangle \\ &= \langle x, S^* T^* y \rangle\end{aligned}$$

So that $(TS)^* = S^* T^*$.

17. Exercise 1.4.17

Recall that any $x \in H$ can be uniquely written as $x = x_V + x_{V^\perp}$. Then, one notes that $\pi_V(x) = x_V$. Hence,

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

So that the adjoint of π_V is precisely the inclusion.

18. Exercise 1.4.18

(i). Note first that

$$||\sum_{i=1}^{N} c_n e_n|| \text{ converges} \iff ||\sum_{i=1}^{N} c_n e_n||^2 \text{ converges}$$

But

$$||\sum_{i=1}^{N} c_n e_n||^2 = \sum_{i=1}^{N} |c_n|^2$$

Letting $N \to \infty$, we see that $\sum_{i=1}^{\infty} |c_n|^2$ must converge.

(ii). Let $\epsilon > 0$. There exists N such that $\sum_{n=N'}^{\infty} |c_n|^2 < \epsilon^2$ for all $N' \geqslant N$. Let $S := \sum_{n=1}^{\infty} c_{\sigma(n)} e_{\sigma(n)}$ be a rearrangement of our sum, and set $M := \max\{\sigma(1), \dots, \sigma(N)\}$. Then, using part (i), we see that for all $M' \geqslant M$:

$$\begin{aligned} ||S - \sum_{m=1}^{M'} c_{\sigma(m)} e_{\sigma(m)} &= || \sum_{m \geqslant M'} c_{\sigma(m)} e_{\sigma(m)} || \\ &= \Big(\sum_{m \geqslant M'} |c_{\sigma(m)}|^2 \Big)^{1/2} \\ &\leqslant \Big(\sum_{n \geqslant N} |c_n|^2 \Big)^{1/2} < \epsilon \end{aligned}$$

Hence, $\sum_{n=1}^{\infty} c_{\sigma(n)} e_{\sigma(n)}$ converges to the same value.

(iii). We see that, given $(a_n), \ (b_n) \in \ell^2(\mathbb{N})$:

$$\begin{aligned} \langle \sum_{n=1}^{\infty} a_n e_n, \sum_{n=1}^{\infty} b_n e_n \rangle &= \sum_{n=1}^{\infty} \langle a_n e_n, b_n e_n \rangle \\ &= \sum_{n=1}^{\infty} a_n \overline{b_n} \langle e_n, e_n \rangle \\ &= \sum_{n=1}^{\infty} a_n \overline{b_n} \\ &= \langle (a_n), (b_n) \rangle \end{aligned}$$

(iv). As already shown, the adjoint of the inclusion is just π_V. Let $x \in H$. Then,

$$\begin{aligned} \langle x, \sum_{n=1}^{\infty} c_n e_n \rangle &= \sum_{n=1}^{\infty} c_n \langle x, e_n \rangle \\ &= \sum_{n=1}^{\infty} c_n \langle x, e_n \rangle \langle e_n, e_n \rangle \\ &= \langle \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n, \sum_{n=1}^{\infty} c_n e_n \rangle \end{aligned}$$

So that

$$\pi_V(x) = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n$$

From part (i), we easily see that

$$||\pi_V(x)||^2 = \sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2$$
$$\implies ||\pi_V(x)|| = \Big(\sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 \Big)^{1/2}$$

And, as $||\pi_V(x)|| \leqslant ||x||$ (by orthogonal decomposition), we see

$$\sum_{n=1}^{\infty} |\langle x, e_n \rangle|^2 \leqslant ||x||^2$$

Which completes the proof.

19. Exercise 1.4.19

$(i) \implies (ii)$: Suppose $(e_\alpha)_{\alpha \in A}$ gives all of H. Then, given $\epsilon > 0$ and $x \in H$, we can find $N \in \mathbb{N}$ such that

$$||x - \sum_{n=1}^{N} c_n e_n|| < \epsilon$$

Implying that finite linear combinations are dense.

$(ii) \implies (iii)$: By Bessel's inequality,

$$\sum_{\alpha \in A} |\langle x, e_\alpha \rangle|^2 \leqslant ||x||^2$$

Note, however, that if V denotes our space of finite linear combinations, $\overline{V} = H \implies V^\perp = \{0\}$. Every vector can be decomposed as $x = x_V + x_{V^\perp}$, where $x_V \perp x_{V^\perp}$. However, this implies that $x_{V^\perp} = 0$, so that

$$||x|| = ||x_V|| \implies \sum_{\alpha \in A} |\langle x, e_\alpha \rangle|^2 = ||x||^2$$

(iii) $\implies$ *(iv)*: We see that the $\langle x, e_\alpha \rangle$ are square summable. By Exercise 1.4.18, part *(ii)*,

$$\sum_{\alpha \in A} \langle x, e_\alpha \rangle e_\alpha$$

converges unconditionally to x.

(iv) $\implies$ *(v)*: Suppose that $x = \sum_{\alpha \in A} \langle x, e_\alpha \rangle e_\alpha$ for all $x \in H$. If $\langle v, e_\alpha \rangle = 0$ for all $\alpha \in A$, then

$$\langle x, v \rangle = 0 \text{ for all } x \in H \iff v = 0$$

(v) $\implies$ *(vi)*: The isomorphism $\ell^2(A) \to H$ is precisely the identification

$$(c_\alpha)_{\alpha \in A} \mapsto \sum_{\alpha \in A} c_\alpha e_\alpha$$

This has already been shown as an isometry, so it merely remains to prove surjectivity. Letting V denote our formal span, we have that $V^\perp = \{0\}$ so that $\overline{V} = H$. But V is a closed set, so in fact $V = H$. One immediately notes that V is precisely the image of our isometry, so surjectivity follows immediately.

(vi) $\implies$ *(i)*: Let T denote our isomorphism. Givne $x \in H$, x is the image of some $(c_\alpha)_{\alpha \in A} \in \ell^2(A)$. If the δ_α denote our standard unit basis vectors, we may rewrite

$$\begin{aligned} (c_\alpha)_{\alpha \in A} &= \sum_{\alpha \in A} c_\alpha (\delta_\alpha)_{\alpha \in A} \\ \implies T(c_\alpha)_{\alpha \in A} &= \sum_{\alpha \in A} c_\alpha e_\alpha \quad \text{(Linearity)} \end{aligned}$$

But, $T(c_\alpha) = x$, so that in fact,

$$x = \sum_{\alpha \in A} c_\alpha e_\alpha$$

Which shows that our Hilbert space span is all of H.

20. Exercise 1.4.20

If V is empty, we are done. Assume $V \neq \varnothing$. Every singleton set is linearly independent, so order the family of linearly independent sets by inclusion. Given any chain $S_1 \subset S_2 \subset \dots$, we have the trivial upper bound $\bigcup_{\lambda \in \Lambda} S_\lambda$. Applying Zorn's Lemma, there exists a maximal linearly independent subset S. It remains to show that $\mathrm{Span}(S) = V$.

Suppose then that $\mathrm{Span}(S) \neq V$. We can choose $v \in V \backslash \mathrm{Span}(S)$, which implies that the set $S \cup \{v\}$ is a linearly independent set that strictly contains S. This contradicts maximality of S, so we conclude that

$$\mathrm{Span}(S) = V$$

21. Exercise 1.4.21

We can assume that A and B are infinite, since the finite case uses the exact same technique without employing the Bernstein Schöder theorem. Let $\{v_\alpha\}_{\alpha \in A}$ and $\{u_\beta\}_{\beta \in B}$ be bases for $\ell^2(A)$ and $\ell^2(B)$, respectively. Then, for each $\alpha \in A$, there exists a finite subset $B_\alpha \subset B$ such that $Tv_\alpha \in \mathrm{Span}\{u_\beta\}_{\beta \in B_\alpha}$. This gives that

$$\mathrm{Span}\{Tv_\alpha\}_{\alpha \in A} \subset \mathrm{Span}\{u_\beta\}_{\beta \in \bigcup_{\alpha \in A} B_\alpha}$$

So that $|A| \leqslant |B|$. However, the same argument applied to $\{T^{-1}u_\beta\}_{\beta \in B}$ shows that $|B| \leqslant |A|$. Applying the Bernstein Schröder theorem, we conclude that $|A| = |B|$.

The converse is trivial, as we merely relabel our indices based off of the provided bijection $f : A \to B$.

22. Exercise 1.4.22

Every basis of a vector space must have the same cardinality, since one can see that if $\text{Span}(A) = \text{Span}(B)$ for two linearly independent sets A, B, that $|A| = |B|$.

To see this, let $\{v_i\}_{i\in I}$, $\{u_j\}_{j\in J}$ be two bases. For each i, there is a finite subset $J_i \subset J_i$ such that

$$v_i \in \text{Span}\{u_j\}_{j\in J_i}$$

Then, we see

$$\text{Span}\{v_i\}_{i\in I} = \text{Span}\{u_j\}_{j\in\bigcup_{i\in I} J_i}$$

So that $|I| \leqslant |J|$. By symmetry, however, we conclude that $|J| \leqslant |I|$ as well, and again we may employ the Bernstein Schröder theorem or cardinal arithmetic to see that $|I| = |J|$.

23. Exercise 1.4.23

If the dimension is countable, our space H is trivially separable by restricting to the rational coefficients. Suppose now that H is separable. There exists some orthonormal basis $\{e_i\}_{i\in I}$. Suppose for sake of contradiction that I is uncountable. Then,

$$||e_i - e_j||^2 = ||e_i||^2 + ||e_j||^2 = 2$$
$$\implies ||e_i - e_j|| = \sqrt{2} \text{ for } i \neq j$$

Consider $B(e_i, 1/2)$. Enumerate our dense set as $\{\alpha_n\}_{n=1}^{\infty}$. For every i, there is at least one unique n such that $\alpha_n \in B(e_i, 1/2)$, by density. Hence, there is a surjection from $\mathbb{N} \to I$. But this forces I to be countable, as desired.

24. Exercise 1.4.23

The map $\otimes$ is the standard tensor product definition. That is, give $H \times H'$ the product space, then modulo the subspaces generated by

$$(rh, h') - (h, rh'), \quad (h_1 + h_2, h') - (h_1, h') - (h_2, h)$$

and so on, and then take the completion.

(i). Linearity is trivial.

(ii). We have basis elements $\{e_i \otimes e'_j\}_{i,j}$. Define

$$\langle e_j \otimes e'_i, e_{j'} \otimes e_{i'} \rangle_{H \otimes H'} := \langle e_j, e_{j'} \rangle_H \langle e'_i, e'_{i'} \rangle_{H'}$$

Extending by linearity yields

$$\langle x \otimes x', y \otimes y' \rangle_{H \otimes H'} := \langle x, y \rangle_H \langle x', y' \rangle_{H'}$$

(iii). Let $x \otimes x' \in H \otimes H'$. Then, $x \otimes x'$ is the limit of some Cauchy sequence $\{x_n \otimes x'_n\}$. Hence, for all $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$||x \otimes x' - x_n \otimes x'_n|| = ||x - x_n|| \cdot ||x' - x'_n|| < \epsilon$$

for $n > N$. Writing x_n, x'_n as finite linear combinations of elements in H, H', respectively, the result follows.

(This can also be done via the indirect identification that a simple tensor $x_1 \otimes x_2$ is such that $x_1 \otimes x_2(x^*) = x^*(x_1)x_2$ for $x^* \in H^*$, so that

$$x_1 \otimes x_2 : H_1^* \to H_2, \quad x_1 \otimes x_2 : H_2^* \to H_1$$

Depending on which dual space your functional belongs to.

25. Exercise 1.4.25

It suffices to show that $\{\psi_n \otimes \phi_m\}_{m,n\in\mathbb{N}}$ is a maximal orthonormal basis of

$$L^2(X \times Y, \mathcal{X} \times \mathcal{Y}, \mu \times \nu)$$

when $\{\psi_n\}_{n\in\mathbb{N}}$ and $\{\phi_m\}_{m\in\mathbb{N}}$ are maximal orthonormal bases of $L^2(X, \mathcal{X}, \mu)$ and $L^2(Y, \mathcal{Y}, \nu)$, respectively. Suppose $f(x,y)$ is orthogonal to every element of $\{\psi_n \otimes \phi_m\}_{m,n\in\mathbb{N}}$. Then, we see

$$\begin{aligned}
&\int_{X\times Y} f(x,y)\overline{\psi_n \otimes \phi_m}(x,y) d\mu \times \eta(x,y) \\
&= \int_X \Big(\int_Y f(x,y)\overline{\phi_m}(y) d\nu(y) \Big) \overline{\psi_n}(x) d\mu(x) = 0 \\
&\implies \int_Y f(x,y)\overline{\phi_m}(y) d\nu(y) = 0 \quad (\psi_n \text{ maximal}) \\
&\implies f(x,y) = 0 \quad (\phi_m \text{ maximal})
\end{aligned}$$

So $\{\psi_n \otimes \phi_m\}_{m,n\in\mathbb{N}}$ is a maximal orthonormal basis. We also see:

$$\begin{aligned}
&\langle f \otimes f', g \otimes g' \rangle_{L^2(X\times Y, \mathcal{X}\times\mathcal{Y}, \mu\times\nu)} \\
&= \int_{X\times Y} f(x)f'(y)\overline{g(x)g'(y)} d\mu \times \nu \\
&= \int_X f(x)\overline{g(x)} d\mu \int_Y f'(y)\overline{g'(y)} d\nu \\
&= \langle f, g \rangle_{L^2(X,\mathcal{X},\mu)} \cdot \langle f', g' \rangle_{L^2(Y,\mathcal{Y},\nu)} \\
&= \langle f \otimes f', g \otimes g' \rangle_{L^2(X,\mathcal{X},\mu)\otimes L^2(Y,\mathcal{Y},\nu)}
\end{aligned}$$

Whence the result.