Exercise 2. Let E be a normed space over the reals. A hyper-plane H is a subspace of E such that the quotient space E/H has dimension 1.

(a) Show that the closure of any subspace of E is again a subspace of E and conclude that a hyper-plane H is either closed or dense in E.

Proof. Consider some vector subspace of E, say H. We recall the definition of the closure of H, denoted $\bar{H}$, as the following:

$$\bar{H} = \{H \cup \text{ the limit points of } H\} \tag{1}$$

To show that $\bar{H}$ is a vector subspace of E then, we need to verify that it is closed under vector addition and scalar multiplication, and moreover that it contains the zero vector of H.

For this, note that the zero vector $0 \in \bar{H}$, as $0 \in H \subseteq \bar{H}$. Thus for $c \in \mathbb{R}$ and $a, b \in \bar{H}$, it is sufficient to show that:

$$ca + b \in \bar{H} \tag{2}$$

For this, simply note that every point in $\bar{H}$ can be expressed as limit of points in H, (either a constant sequence or some non-trivial sequence for those points in $\bar{H} \setminus H$).

With this in mind, we represent a as $\lim\limits_{n\to\infty} a_n$ and b as $\lim\limits_{n\to\infty} b_n$ respectively, where $\{a_n\}_{n=1}^{\infty} \subseteq H$ and $\{b_n\}_{n=1}^{\infty} \subseteq H$. Using limit rules, we get the following equivalent convergent sequences:

$$ca + b = c \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n = \lim_{n\to\infty} ca_n + b_n \tag{3}$$

But then of course the last sequence in (3) is a sequence of points in H that converges to $ca + b$, so it must be $ca + b \in \bar{H}$. Thus $\bar{H}$ is a vector subspace of E.

We want to conclude then that a hyper-plane H is either closed or dense in E. For this, assume H is not closed. We would like to show H is dense in E, which is tantamount to showing $E \subseteq \bar{H}$.

For this, consider $y \in E$. As H is not closed, we have:

$$\exists \{x_n\}_{n=1}^{\infty} \subseteq H \text{ such that } x_n \to \hat{x} \notin H, \text{ i.e. } \hat{x} \in \bar{H} \setminus H$$

Then as $\hat{x} \notin H$, the coset $\hat{x} + H$ is not the zero vector in the quotient space. As E/H is 1-dimensional then, it is spanned by $\hat{x} + H$.

In particular, we have for some $c \in \mathbb{R}$ that:

$$y + H = c(\hat{x} + H) = c\hat{x} + H$$

Thus $y - c\hat{x} = h \in H \subseteq \bar{H}$, so we can reorganize to get $y = c\hat{x} + h$. But as $\bar{H}$ is a vector subspace,

$$c\hat{x} + h = y \in \bar{H} \text{ given that } c\hat{x}, h \in \bar{H}$$

Thus $E \subseteq \bar{H}$, where we already have $\bar{H} \subseteq E$. Therefore $E = \bar{H}$, which is exactly that H is dense in E.

Therefore H is either closed or dense in E. $\square$

(b) Let u be a linear functional on E. Prove that u is discontinuous if and only if there exists a sequence $\{x_n\}_{n=1}^{\infty} \subseteq E$ that conveges to 0 and for which $u(x_n) = 1$, $\forall n \in \mathbb{N}$.

Proof. We will show the forward implication first. For this, we assume u is discontinuous. As u is a linear functional then, it must be discontinuous at 0, i.e. we have a sequence $\{y_n\}_{n=1}^{\infty} \subseteq E$ such that $y_n \to 0$ but $u(y_n)$ does not converge to 0.

As $u(y_n)$ does not converge to zero, we have for some $\epsilon > 0$ that there is a subsequence $\{y_{n_k}\}_{k=1}^{\infty} \subseteq E$ such that $|u(y_{n_k})| \geq \epsilon$ for all $k \in \mathbb{N}$. Define then a sequence $\{x_k\}_{k=1}^{\infty} \subseteq E$ by stipulating $x_k = \frac{y_{n_k}}{u(y_{n_k})}$. We quickly verify $x \to 0$:

$$||x_k|| = \left|\left|\frac{y_{n_k}}{u(y_{n_k})}\right|\right| = \frac{1}{|u(y_{n_k})|}||y_{n_k}|| \leq \frac{1}{\epsilon}||y_{n_k}|| \tag{4}$$

Of course then, we can take $||y_{n_k}||$ small (as the subsequence still converges to 0), which shows $x_k \to 0$. We want to verify now that $u(x_k) = 1$ for all $k \in \mathbb{N}$. Here, we note:

$$u(x_k) = u\left(\frac{y_{n_k}}{u(y_{n_k})}\right) = \frac{1}{u(y_{n_k})}u(y_{n_k}) = 1 \tag{5}$$

So x_k is the desired sequence, and so the forward implication holds.

We note then that the reverse implication is obvious, as the negation of continuity at 0 is that there is some sequence $\{x_n\}_{n=1}^{\infty} \subseteq E$ such that $x_n \to 0$ but $u(x_n)$ does not converge to $u(0) = 0$.

If we assume then that we have some sequence such that $x_n \to 0$ but $u(x_n) = 1$, $\forall n \in \mathbb{N}$, it is clear that such a sequence fulfills this definition of the negation of continuity, and so u would be discontinuous. $\square$

(c) Let $x_0 \in E$ be a unit norm vector, and H the complement of the one dimensional space spanned by x_0. Show that every $x \in E$ can be uniquely decomposed as:

$$x = t(x)x_0 + y(x) \tag{6}$$

where t and y are linear maps from E to $\mathbb{R}$ and H respectively. Prove that t and y are continuous iff H is closed.

Proof. We adopt the notation $\langle x_0 \rangle = \text{span}(x_0)$. As H is selected to be the complement of $\langle x_0 \rangle$, we have that:

$$E = \langle x_0 \rangle \oplus H$$

From this it follows we can uniquely write any element $x \in E$ as:

$$x = \lambda x_0 + h$$

for some $\lambda \in \mathbb{R}, h \in H$. We denote these elements λ_x and h_x, i.e. $x = \lambda_x x_0 + h_x$. With this in mind, we define our functions $t(x) = \lambda_x$ and $y(x) = h_x$.

(1) *(Decomposition and well-definedness)* By construction, we then get a decomposition $x = t(x)x_0 + y(x)$ for all x.

Moreover, note these functions are well-defined, as the decomposition of x into a sum of elements in complementary subspaces is unique.

(2) *(Linearity)* Linearity also follows almost immediately from the uniqueness of our decomposition. In particular, we note for $x, y \in E$, we of course can write

$$x = \lambda_x x_0 + h_x \text{ and } y = \lambda_y x_0 + h_y$$

Moreover, we can express $x + y = \lambda_{x+y} x_0 + h_{x+y}$, for some $\lambda_{x+y} \in \mathbb{R}$ and $h_{x_y} \in H$.

However, we also have that:

$$x + y = \lambda_x x_0 + h_x + \lambda_y x_0 + h_y = (\lambda_x + \lambda_y)x_0 + (h_x + h_y)$$

Uniqueness of this decomposition gives us then that:

$$\lambda_{x+y} = \lambda_x + \lambda_y \text{ and } h_{x+y} = h_x + h_y$$

Thus $t(x + y) = t(x) + t(y)$ and $y(x + y) = y(x) + y(y)$, and so both functions are linear.

(3) *(Uniqueness)* We finally remark that these functions t, y are necesarily unique, again using the uniqueness of the decomposition into a sum.

In particular, we know we have the unique decomposition $x = \lambda_x x_0 + h_x$, so if we have functions:

$$a : E \to \mathbb{R} \text{ and } b : E \to H \text{ where } x = a(x)x_0 + b(x)$$

as $a(x)x_0 \in \langle x_0 \rangle$ and $b(x) \in H$, the only possibility is that $a(x) = \lambda_x$ and $b(x) = h_x$, which yields the given functions t and y.

(4) *(Continuity of both implies closedness of H)* We want now to show that t and y being continuous implies the closedness of H.

For this, we will consider the contrapositive, that is H being not closed implies that t or y is not continuous. In particular, we will show that t is not continuous if H is not closed.

For this, assume that we have a sequence:

$$\{x_n\}_{n=1}^{\infty} \subseteq H \text{ such that } x_n \to x \notin H$$

We note as $x \notin H$, it must be then that $\lambda_x \neq 0$, i.e. $t(x) \neq 0$. However, as each $x_n \in H$, we know $t(x_n) = \lambda_{x_n} = 0$. Thus:

$$\lim_{n\to\infty} t(x_n) = 0 \neq 1 = t(x)$$

So t is not continuous at x, the desired result.

(5) *(Closedness of H implies continuity of both t and y)* Note it is enough to prove that one is continuous, i.e. t is continuous, because we can reorganize to get $y(x) = x - t(x)x_0$.

In particular, $t : E \to \mathbb{R}$ being continuous has $tx_0 : E \to E$ being continuous, which has y continuous as the difference of the identity and this function.

So we want to show H being closed has t continuous. For this, we assume for the sake of contradiction that t is discontinuous.

Then as t is a linear functional on E, we apply our result in (b) to get some sequence:

$$\{x_n\}_{n=1}^{\infty} \subseteq E \text{ such that } x_n \to 0 \text{ but } t(x_n) = 1$$

for all $n \in \mathbb{N}$.

Then we get $y(x_n) = x_n - x_0$. So taking the limit gets $\lim\limits_{n\to\infty} y(x_n) = -x_0$. But this is an issue, as y maps into H, so $\{y(x_n)\}_{n=1}^{\infty} \subseteq H$.

As H is closed then, it should be $-x_0 \in H$, but of course $-x_0 \neq 0 \in \langle x_0 \rangle$ and so as H is a complementary subspace, it must be then that $-x_0 \notin H$.

So we have a contradiction, and thus H being closed has t and furthermore y continuous.

□

(d) Let u be a linear functional on E. Prove that u is continuous if and only if H, the kernel, is closed.

Proof. We will show both directions seperately.

(1) *(Continuity of u implies the kernel is closed)* Say u is continuous, and assume for contradiction that $H = \ker(u)$ is not closed.

By definition then, there exists some sequence

$$\{x_n\}_{n=1}^{\infty} \subseteq \ker(u) \text{ such that } x_n \to x \notin \ker(u), \text{ i.e. } u(x) \neq 0$$

We note additionally then that for any $n \in \mathbb{N}$, we have that $u(x_n) = 0$, as each term in in the kernel.

We can apply continuity then:

$$0 = \lim_{n\to\infty} u(x_n) = u(\lim_{n\to\infty} x_n) = u(x) \neq 0 \tag{7}$$

But of course this generates a contradiction. So it must be $\ker(u)$ is closed.

(2) *(The closedness of the kernel implies the continuity of u)* Say that $\ker(u)$ is closed, and assume for contradiction that u is not continuous.

As u is not continuous, it is not bounded, i.e. for $\forall c > 0$, there exists $x \in E$ such that:

$$|u(x)| > c||x||$$

Consider the sequence $\{x_n\}_{n=1}^{\infty}$ generated by taking $c = n$ in the above definition, i.e. $|u(x_n)| \geq n||x_n||$.

Take then some point $x \notin \ker(u)$ (if no such point exists, then u is identically zero and thus already continuous). We define the following sequence:

$$\hat{x}_n = x - \frac{u(x)}{u(x_n)} x_n \tag{8}$$

Note then:

$$\begin{aligned} \hat{x}_n = u\left(x - \frac{u(x)}{u(x_n)} x_n\right) &= u(x) - u\left(\frac{u(x)}{u(x_n)} x_n\right) \\ &= u(x) - \frac{u(x)}{u(x_n)} u(x_n) = u(x) - u(x) = 0 \end{aligned} \tag{9}$$

Thus $\{\hat{x}_n\}_{n=1}^{\infty} \subseteq \ker(u)$. We claim then that $\frac{u(x)}{u(x_n)} x_n \to 0$ as $n \to \infty$. For this we note the following argument:

$$\begin{aligned} \left|\left|\frac{u(x)}{u(x_n)} x_n\right|\right| = \left|\frac{u(x)}{u(x_n)}\right| ||x_n|| &= \frac{|u(x)|}{|u(x_n)|} ||x_n|| \\ &= \frac{||x_n||}{|u(x_n)|} |u(x)| < \frac{|u(x)|}{n} \end{aligned} \tag{10}$$

Taking n large thus demonstrates that $\frac{u(x)}{u(x_n)} x_n \to 0$ as $n \to \infty$.

Thus $\lim_{n\to\infty} \hat{x}_n = x \notin \ker(u)$, but this is a contradiction as $\{\hat{x}_n\}_{n=1}^{\infty} \subseteq \ker(u)$ and $\ker(u)$ is closed. So it must be u is continuous.

$\square$

Exercise 5. Let E be a Banach space.

(a) Prove that if $T \in L(E,E)$ and $||I - T|| < 1$ where I is the identity operator, then T is invertible, and that the series $\sum_{n=0}^{\infty}(I-T)^n$ converges in $L(E,E)$ to T^{-1}.

Proof. For simplicitly, we denote $S = I - T$, and observe clearly $S \in L(E,E)$. As E is complete with respect to its norm metric, so is $L(E,E)$ by a theorem in Folland.

Note then, using submultiplicity of the norm metric, we get the following:

$$\sum_{n=0}^{\infty} ||S^n|| \leq \sum_{n=0}^{\infty} ||S||^n < \infty \tag{11}$$

In particular, convergence here follows from the geometric series test, as $||S|| < 1$.

As $L(E,E)$ is complete then, we have that the series $\sum_{n=0}^{\infty} S^n$ converges to a $T^{-1} \in L(E,E)$ (we are not assuming this is the inverse of T, just calling it this for now).

We want then to actually verify this $T^{-1} = \sum_{n=0}^{\infty} S^n \in L(E,E)$ is the inverse of T. For this, we need to show composition on both sides yields the identity operator I.

We start by verifying it is a right inverse, using the linearity of T:

$$\begin{aligned} T\sum_{n=0}^{\infty} S^n = T\lim_{k\to\infty}\sum_{n=0}^{k} S^n = \lim_{k\to\infty}\sum_{n=0}^{k} TS^n = \lim_{k\to\infty}\sum_{n=0}^{k}(I-S)S^n \\ = \lim_{k\to\infty}\sum_{n=0}^{k} S^n - S^{n+1} = \lim_{k\to\infty} I - S^{k+1} \end{aligned} \tag{12}$$

Using the divergence test on the first series in (11), we note that it is that $\lim_{k\to\infty}||S^k|| = 0$, which is that $S^k \xrightarrow{L(E,E)} 0$, where 0 is the zero operator.

Thus the last limit in (12) is just the identity, and thus $\sum_{n=0}^{\infty} S^n$ is a right inverse.

For verifying that it also left inverse, we note the same argument works as $(I-S)S^n = S^n(I-S)$, and so the argument ends being the exact same.

Thus T is invertible with $T^{-1} = \sum_{n=0}^{\infty} S^n \in L(E,E)$, and of course by construction our series $\sum_{n=0}^{\infty} S^n \in L(E,E)$ converges to T^{-1}. $\square$

(b) Show that if $T \in L(E,E)$ is invertible and $||S-T|| < ||T^{-1}||^{-1}$, then S is invertible. Conclude that the set of invertible operators in $L(E,E)$ is open.

Proof. We get the following use submultiplicity of the operator norm:

$$||T^{-1}S - I|| = ||T^{-1}(S-T)|| \leq ||T^{-1}||\,||S-T|| < ||T^{-1}||\,||T^{-1}||^{-1} = 1 \tag{13}$$

Using part (a) then, it follows that $T^{-1}S$ is invertible, i.e. there is some $R \in L(E,E)$ such that $RT^{-1}S = T^{-1}SR = I$.

Of course, this shows RT^{-1} is a left inverse of S. Given $T^{-1}SR = I$ we can reorganize:

$$T^{-1}SR = I \to SR = T \to SRT^{-1} = I \tag{14}$$

So S is invertible with inverse RT^{-1}.

For the conclusion that the set of invertible operators is open in $L(E,E)$, we just note that this has for an invertible operator T that $S \in B_{||T^{-1}||^{-1}}(T)$ is also invertible, i.e.:

$$B_{||T^{-1}||^{-1}}(T) \subseteq \{\text{invertible operators in } L(E,E)\}$$

but this is precisely the definition of openness is the norm topology. □

Exercise 8. Suppose that H is a Hilbert space and $T \in L(H,H)$.

(a) Prove that there exists a unique $T^* \in L(H,H)$, called the adjoint of T, such that $\langle Tx, y\rangle = \langle x, T^*y\rangle$, for all $x, y \in H$.

Proof. Let $y \in H$, then $\langle Tx, y\rangle$ is a functional in x. Linearity of this functional immediately follows from the linearity of the inner product and T.

We can furthermore see it is bounded by the Cauchy-Schwarz inequality (and that T is bounded, i.e. $||Tx|| \leq c||x||$ for $c > 0$):

$$|\langle Tx, y\rangle| \leq ||Tx||\,||y|| \leq (c||y||)||x|| \tag{15}$$

And so it is bounded. As it a linear, bounded functional then, we can apply the **Riesz Representation Theorem** to get a unique $z \in H$ for which we have the following:

$$\langle Tx, y\rangle = \langle x, z\rangle, \forall x \in H \tag{16}$$

With this in mind, we define $T^*y = z$. It is clear this mapping is unique, as z for which (16) is satisfied is unique. Moreover, by construction, it is that $\langle Tx, y\rangle = \langle x, T^*y\rangle$ for all $x, y \in H$.

We need to verify this definition yields an operator that is linear and bounded. We start with linearity, we would like to show $T^*(ag + bh) = aT^*g + bT^*h$ where $g, h \in H$

and $a, b \in \mathbb{C}$.

For this, we note the following argument:

$$\begin{aligned}
&\langle x, T^*(ag+bh)\rangle \\
&= \langle Tx, ag+bh\rangle \\
&= \bar{a}\langle Tx, g\rangle + \bar{b}\langle Tx, h\rangle \\
&= \bar{a}\langle x, T^*g\rangle + \bar{b}\langle x, T^*h\rangle \\
&= \langle x, aT^*g\rangle + \langle x, bT^*h\rangle \\
&= \langle x, aT^*g + bT^*h\rangle
\end{aligned} \tag{17}$$

Thus:

$$\langle x, T^*(ag+bh)\rangle = \langle Tx, ag+bh\rangle = \langle x, aT^*g + bT^*h\rangle, \forall x \in H \tag{18}$$

But as we've already shown $\langle Tx, ag+bh\rangle$ is a linear bounded functional, and so it is equivalent to $\langle x, z\rangle$ for all x for some unique z.

Uniqueness of this z thus tells us from (18) it is that $T^*(ag+bh) = aT^*g + bT^*h$, which is precisely linearity.

We need finally to show boundedness then. Here, will we again appeal to the **Riesz Representation Theorem**. Recall that by this theorem and the definition of T^*y, we have that $||T^*y|| = ||\langle Tx, y\rangle||$, where the latter term is the operator norm.

In particular, we get the following:

$$\begin{aligned}
||T^*y|| = ||\langle Tx, y\rangle|| &= \sup_{||x||=1} |\langle Tx, y\rangle| \\
&\leq \sup_{||x||=1} ||Tx||\,||y|| = ||y||\,||T||
\end{aligned} \tag{19}$$

So T^* is bounded with $c = ||T||$, and so we finally have $T^* \in L(H, H)$.

Putting this all together has that we have a unique T^* such that $\langle Tx, y\rangle = \langle x, T^*y\rangle$, for all $x, y \in H$, and moreover $T^* \in L(H, H)$. $\square$

(b) Prove that $T^* = V^{-1}T^\dagger V$, where V is the conjugate-linear isomorphism from H to H^* given by $(Vy)(x) = \langle x, y\rangle$.

Proof. For $T \in L(H, H)$, we have $\langle x, T^*y\rangle = \langle Tx, y\rangle$, for all $x, y \in H$. Using this, we can get the following:

$$\begin{aligned}
&\langle x, T^*y\rangle = \langle Tx, y\rangle \\
&\to (VT^*y)(x) = \langle Tx, y\rangle
\end{aligned} \tag{20}$$

Moreover, we have the following:

$$\langle Tx, y\rangle = (Vy)(Tx) = (Vy \circ T)(x) = T^\dagger Vy(x) \tag{21}$$

We use this to continue (20):

$$\begin{aligned}
(VT^*y)(x) &= \langle Tx, y\rangle \\
\to (VT^*y)(x) &= T^\dagger Vy(x), \text{ still for all } x \in H \\
\to VT^*y &= T^\dagger Vy, \text{ still for all } y \in H \\
\to VT^* &= T^\dagger V \\
\to T^* &= V^{-1}T^\dagger V
\end{aligned} \tag{22}$$

Which is the desired result. □

(c) Prove that:

(i) $||T^*|| = ||T||$

Proof. We will establish this by showing that $V : H \to H^*$ and $V^{-1} : H^* \to H$ as given in the previous part both have operator norm 1.

We first move to compute the norm of V:

$$||V|| = \sup_{||y||=1} ||Vy|| = \sup_{||y||=1} \sup_{||x||=1} |\langle x, y\rangle| \tag{23}$$

On one hand:

$$\sup_{||y||=1} \sup_{||x||=1} |\langle x, y\rangle| \leq \sup_{||y||=1} \sup_{||x||=1} ||x||\,||y|| = 1 \tag{24}$$

But then also clearly this quantity $||V||$ must bound above $|\langle x, x\rangle|$ for some x where $||x|| = 1$, i.e. $||V|| \geq |\langle x, x\rangle| = 1$. Thus $||V|| = 1$.

To ascertain the value of $||V^{-1}||$, we first need to explicitly write out what V^{-1} is.

For this, we claim that $V^{-1} : H^* \to H$ is the mapping given by $f \mapsto z$, where f is a bounded linear functional and z is the unique point associated with it by the **Riesz Representation Theorem**.

We verify this is indeed the inverse:

$$\begin{aligned}
V^{-1}Vy = V^{-1}\langle\cdot, y\rangle = y \\
VV^{-1}f = Vz = \langle\cdot, z\rangle = f
\end{aligned} \tag{25}$$

Note here we are making extensive use the **Riesz Representation Theorem** (in particular, the uniqueness of z).

Thus the described mapping is V^{-1}, and we can ascertain its norm again by using the **Riesz Representation Theorem**, in particular using the fact that:

$$||f|| = ||\text{the } z \text{ associated with it}||$$

We consequently have the following:

$$||V^{-1}|| = \sup_{||f||=1} ||V^{-1}f|| = \sup_{||f||=1} ||z|| = \sup_{||f||=1} ||f|| = 1 \tag{26}$$

And thus $||V^{-1}|| = 1$. Using the previous part and some reorganizing then, we have that both $T^* = V^{-1}T^\dagger V$ and $VT^*V^{-1} = T^\dagger$. Recalling from **Problem 6** that $||T^\dagger|| = ||T||$ then, we get the following using submultiplicity:

$$\begin{aligned} ||T^*|| = ||V^{-1}T^\dagger V|| \le ||V^{-1}||\,||T^\dagger||\,||V|| = ||T^\dagger|| = ||T|| \\ ||T|| = ||T^\dagger|| = ||VT^*V^{-1}|| \le ||V||\,||T^*||\,||V^{-1}|| = ||T^*|| \end{aligned} \tag{27}$$

Thus $||T|| = ||T^*||$. ☐

(ii) $||TT^*|| = ||T||^2$

Proof. Obviously, from the prior part, we get $||TT^*|| \le ||T||^2$, as this is just:

$$||TT^*|| \le ||T||\,||T^*|| = ||T||^2 \tag{28}$$

We want to show $||T||^2 \le ||TT^*||$. For this, we note the following argument:

$$\begin{aligned} ||T||^2 &= \Big(\sup_{||x||=1} ||Tx||\Big)^2 \\ &= \sup_{||x||=1} ||Tx||^2 = \sup_{||x||=1} |\langle Tx, Tx\rangle| \\ &= \sup_{||x||=1} |\langle x, T^*Tx\rangle| \le \sup_{||x||=1} ||x||\,||T^*Tx|| \\ &= \sup_{||x||=1} ||T^*Tx|| = ||T^*T|| \end{aligned} \tag{29}$$

Thus $||TT^*|| = ||T||^2$. ☐

(iii) $(aS + bT)^* = \bar{a}S * + \bar{b}T^*$

Proof. For this, we note the following argument using the sesquilinear properties of the inner product:

$$\begin{aligned} \langle x, (aS + bT)^*y\rangle &= \langle (aS + bT)x, y\rangle \\ &= \langle aSx + bTx, y\rangle = a\langle Sx, y\rangle + b\langle Tx, y\rangle \\ &= a\langle x, S^*y\rangle + b\langle x, T^*y\rangle = \langle x, \bar{a}S^*y\rangle + \langle x, \bar{b}T^*y\rangle \\ &= \langle x, \bar{a}S^*y + \bar{b}T^*y\rangle, \forall x, y \in H \end{aligned} \tag{30}$$

The idea given after (17) concludes then that it must be $(aS + bT)^* = \bar{a}S * + \bar{b}T^*$. However, we can also see this in a more general way by the following, fixing a $y \in H$:

$$\begin{aligned} \langle x, (aS + bT)^*y\rangle &= \langle x, \bar{a}S^*y + \bar{b}T^*y\rangle \\ \to \langle x, (aS + bT)^*y\rangle - \langle x, \bar{a}S^*y + \bar{b}T^*y\rangle &= 0 \\ \to \langle x, (aS + bT)^*y\rangle + \langle x, -\bar{a}S^*y - \bar{b}T^*y\rangle &= 0 \\ \to \langle x, (aS + bT)^*y - \bar{a}S^*y - \bar{b}T^*y\rangle &= 0, \forall x \in H \end{aligned} \tag{31}$$

The only vector orthogonal to all others is the zero vector then, so $(aS+bT)^*y - \bar{a}S^*y - \bar{b}T^*y = 0$, $\forall y \in H$, i.e.:

$$\begin{aligned}(aS+bT)^*y &= \bar{a}S^*y + \bar{b}T^*y \\ \rightarrow (aS+bT)^*y &= (\bar{a}S^* + \bar{b}T^*)y, \forall y \in H \\ \rightarrow (aS+bT)^* &= \bar{a}S^* + \bar{b}T^*\end{aligned} \tag{32}$$

I will use this same motif in later parts by appealing to it here. □

(iv) $(ST)^* = T^*S^*$ We use the same type of argument as the previous parts:

Proof.

$$\begin{aligned}\langle x, (ST)^*y\rangle &= \langle STx, y\rangle \\ = \langle Tx, S^*y\rangle &= \langle x, T^*S^*y\rangle\end{aligned} \tag{33}$$

The same trick as the previous part gets $(ST)^* = T^*S^*$. □

(v) $T^{**} = T$ Again, we use a similar argument:

Proof.

$$\begin{aligned}\langle x, T^{**}y\rangle &= \langle T^*x, y\rangle \\ = \overline{\langle y, T^*x\rangle} &= \overline{\langle Ty, x\rangle} \\ &= \langle x, Ty\rangle\end{aligned} \tag{34}$$

The same idea again concludes $T^{**} = T$. □

(d) Let $R(T)$ and $N(T)$ denote the range and nullspace of T respectively. Prove that $R(T)^\perp = N(T^*)$ and $N(T)^\perp = \overline{R(T^*)}$.

Proof. We will first show that $R(T)^\perp = N(T^*)$. For this, let $x \in R(T)^\perp$. This is precisely that for $\forall y \in H$ we have $\langle x, Ty\rangle = 0$. Using that $T^{**} = T$ then, this is that:

$$0 = \langle x, Ty\rangle = \langle x, T^{**}y\rangle = \langle T^*x, y\rangle, \forall y \in H$$

Again, the only vector orthogonal to all other vectors is the the zero vector. Thus $T^*x = 0$, and so $x \in N(T^*)$. It follows $R(T)^\perp \subseteq N(T^*)$.

Consider $x \in N(T^*)$, then $T^*x = 0$. Thus $\langle T^*x, y\rangle = 0, \forall y \in H$. Continuing this we get (again using $T^{**} = T$):

$$0 = \langle T^*x, y\rangle = \langle x, Ty\rangle, \forall y \in H$$

As every term in $R(T)$ can be expressed as Ty then for some $y \in H$, it follows that $x \in R(T)^\perp$. Thus $R(T)^\perp \subseteq N(T^*)$, and so $R(T)^\perp = N(T^*)$.

We want then to prove that $N(T)^\perp = \overline{R(T^*)}$. Note first, that by substituting T^* in our previous work, we get $R(T^*)^\perp = N(T)$ (we can do this as $T^* \in L(H,H)$ and $T^{**} = T$).

Thus $\left(R(T^*)^\perp\right)^\perp = N(T)^\perp$. Recall then $\left(R(T^*)^\perp\right)^\perp = \overline{\text{span}R(T^*)}$.

Obviously $R(T^*) \subseteq \text{span}R(T^*)$. Consider then some $x \in \text{span}R(T^*)$, then

$$x = \sum_{n=1}^{N} c_n T^* y_n = T^*\Big(\sum_{n=1}^{N} c_n y_n\Big)$$

using the linearity of T^*. But then of course $\sum_{n=1}^{\infty} c_n y_n \in H$, so it is that x is in the range of T^*, i.e. $x \in R(T*)$. Thus $R(T^*) = \text{span}R(T^*)$.

It follows $\overline{R(T^*)} = \left(R(T^*)^\perp\right)^\perp = N(T)^\perp$, which gives the desired result. □

(e) Show that T is unitary if and only if T is invertible and $T^{-1} = T^*$.

Proof. Assume T is unitary, then we have it is surjective and $\langle Tx, Ty\rangle = \langle x, y\rangle, \forall x, y \in H$. Thus:

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

Consider some $y \in H$. Using this we get the following argument:

$$\begin{aligned} \langle x, T^*Ty\rangle &= \langle x, y\rangle \\ \to \langle x, T^*Ty\rangle - \langle x, y\rangle &= 0 \\ \to \langle x, T^*Ty\rangle + \langle x, -y\rangle &= 0 \\ \to \langle x, T^*Ty - y\rangle &= 0, \forall x \in H \end{aligned} \tag{35}$$

Using the same argument as earlier, thus $T^*Ty - y = 0 \to T^*Ty = y$, $\forall y \in H$, i.e. $T^*T = I$.

T has a left inverse then, and so it is injective. This thus has it is a bijection, and so it admits a unique inverse T^{-1}.

Furthermore, we note $T^*TT^{-1} = T^{-1} \to T^* = T^{-1}$. So $T^* \in L(H, H)$ is the inverse of T, and so T is invertible with $T^{-1} = T^*$.

Say then for the reverse implication that T is invertible and $T^{-1} = T^*$. Obviously, this has that T is surjective. We want to show then that the inner product is preserved, i.e.:

$$\langle Tx, Ty\rangle = \langle x, y\rangle,\ \forall x, y \in H \tag{36}$$

For this, we just note the following argument:

$$\langle Tx, Ty\rangle = \langle x, T^*Ty\rangle = \langle x, y\rangle \tag{37}$$

It follows T is unitary. □

Exercise 12. Let M be a closed subspace of $L^2([0,1])$ that is contained in $C([0,1])$.

(a) Prove that there exists a $C > 0$ such that $||f||_u \le C||f||_2$ for all $f \in M$.

Proof. Consider the space $(M, ||\cdot||_u)$, we would like to show it is a closed subspace of $(C[0,1], ||\cdot||_u)$.

We first note that $C[0,1]$ is a subspace of $L^2[0,1]$ (as the continuous functions are closed under sums and scaling).

As M is also a subspace of $L^2([0,1])$ then and $C[0,1]$ contains M, it naturally follows M is a subspace of $C[0,1]$.

We want then to show M is closed in $C[0,1]$ under the topology generated by the supremum norm.

By assumption, M is closed under the norm topology generated by the L^2 norm, i.e.:

$$\text{if } \{f_n\}_{n=1}^{\infty} \subseteq M \text{ and } f_n \xrightarrow{L_2} f \in L^2[(0,1)] \text{ then } f \in M$$

Consider then some sequence of functions $\{f_n\}_{n=1}^{\infty} \subseteq M$ such that:

$$f_n \to f \in C[0,1] \subseteq L^2([0,1]) \text{ uniformly}$$

We recall uniform convergence implies convergence in L^2, so using the closedness of M with respect to the L^2 norm in $L^2([0,1])$ thus implies that $f \in M$.

So M is a closed subset of $C[0,1]$ under the topology induced by the supremum norm.

It follows $(M, ||\cdot||_u)$ is a closed subspace of $(C[0,1], ||\cdot||_u)$. As the latter is Banach then (a well-known result), so is the former (as closed subspaces of Banach spaces are also Banach spaces).

Similarly, this has that $(M, ||\cdot||_2)$ is a Banach space. With this in mind, we define a mapping I by the following specifications:

$$\begin{aligned} I : (M, ||\cdot||_u) &\to (M, ||\cdot||_2) \\ I : f &\mapsto f \end{aligned} \tag{38}$$

Trivially, this map is linear, as it is essentially the identity map. Moreover, we remark it is continuous as uniform convergence implies convergence in the L^2 norm.

In particular, that is whenever we have a sequence of functions $\{f_n\}_{n=1}^{\infty}$ such that $f_n \to f$ uniformly, we naturally have $I(f_n) = f_n \xrightarrow{L_2} f = I(f)$.

As I is continuous then, it is bounded. Moreover, it is trivially a bijection. Appealing to the **Bounded Inverse Theorem** then, it follows that I has a bounded inverse.

Of course this inverse I^{-1} is also the identity mapping, but we now know it is bounded. In particular, this means for some $C > 0$:

$$||I^{-1}f||_u = ||f||_u \leq C||f||_2 \tag{39}$$

Which is the desired result. $\square$

(b) For each $x \in [0,1]$, prove that there exists a $g_x \in M$ such that $f(x) = \langle f, g_x \rangle$ for all $f \in M$ and that $||g_x||_2 \leq C$.

Proof. Let some $x \in [0,1]$. Consider the functional $T_x : (M, ||\cdot||_2) \to \mathbb{C}$ defined by $T_x : f \mapsto f(x)$.

We quickly verify T_x is linear:

$$\begin{aligned} T_x(af + bg) &= (af + bg)(x) \\ = af(x) + bg(x) &= aT_x(f) + bT_x(g) \end{aligned} \tag{40}$$

Moreover, it is clearly bounded as:

$$|f(x)| \leq ||f||_u \leq C||f||_2 \tag{41}$$

per the previous part. Note then that as M is a closed subspace of $L^2([0,1])$, we have that $(M, ||\cdot||_2)$ is a Hilbert space under the inherited inner product structure.

Appealing to the **Riesz Representation Theorem** then, we have that there is some unique $g_x \in M$ such that:

$$f(x) = T_x(f) = \langle f, g_x \rangle,\ \forall f \in M \tag{42}$$

Moreover, we have that:

$$||g_x||_2 = ||T_x|| = \sup_{||f||_2 = 1} |f(x)| \tag{43}$$

But when $||f||_2 = 1$, we naturally have:

$$|f(x)| \leq ||f||_u \leq C||f||_2 = C \tag{44}$$

So as we have here that$|f(x)| \leq C$, it follows:

$$||g_x||_2 = ||T_x|| = \sup_{||f||_2 = 1} |f(x)| \leq C \tag{45}$$

Thus our g_x satisifes the desired properties. $\square$

(c) Show that the dimension of M is at most C^2 by proving that if $\{f_k\}_{k=1}^{\infty}$ is any orthonormal sequence in M then $\sum_{k=1}^{\infty} |f_k(x)|^2 \leq C^2$ for all $x \in [0,1]$.

Proof. Let $x \in [0,1]$ Using our work in part (b) (i.e. generating a g_x), this becomes a routine application of **Bessel's inequality**, which applies as we are given an orthonormal sequence.

$$\begin{aligned}\sum_{k=1}^{\infty} |f_k(x)|^2 &= \sum_{k=1}^{\infty} |\langle f_k, g_x \rangle| \\ &= \sum_{k=1}^{\infty} |\langle g_x, f_k \rangle| \leq ||g_x||_2^2 \leq C^2\end{aligned} \tag{46}$$

Note we are able to consider this sequence as countable, as the M inherits the seperability of L^2 (given it is a closed subspace), and so every orthnormal sequence in M is countable.

Consider then an ONB for M given by $\{f_k\}_{k=1}^{\infty}$ (all Hilbert spaces admit orthonormal bases). We get the following:

$$\sum_{k=1}^{\infty} \langle f_k, f_k \rangle = \sum_{k=1}^{\infty} ||f_k||_2^2 = \sum_{k=1}^{\infty} \int_0^1 |f_k|^2 \tag{47}$$

$$= \int_0^1 \sum_{k=1}^{\infty} |f_k|^2 \qquad \text{(Swapping sum and integral, as } |f_k| > 0) \tag{48}$$

$$\leq \int_0^1 C^2 = C^2 \qquad \text{(Using previous work)} \tag{49}$$

But of course $\sum_{k=1}^{\infty} \langle f_k, f_k \rangle$ should be the amount of nonzero terms in $\{f_k\}_{k=1}^{\infty}$, as its orthonormal.

Therefore the amount of nonzero terms is bounded above by C^2, i.e. we have a basis for M with at most C^2 vectors. Thus the dimension of M is at most C^2. □

Exercise 20. Suppose $||f_0||_p = ||f_1||_p = 1$ and let:

$$f_t = (1-t)f_0 + f_1 \tag{50}$$

be the straight line segment joining the points f_0 and f_1. Then $||f_t||_p < 1$ for all t with $0 < t < 1$, unless $f_0 = f_1$.

(a) Let $f \in L^p$ and $g \in L^q$, p, q dual, with $||f||_p = 1$ and $||g||_q = 1$. Then:

$$\int fg d\mu = 1 \tag{51}$$

only when $f(x) = \text{sign}|g(x)|^{q-1}$.

Proof. We first verify that taking $f(x) = \text{sign}g(x)|g(x)|^{q-1}$ has that $\int fg d\mu = 1$, which we can do simply by noting $|g| = \text{sign}g \cdot g$.

$$\begin{aligned} \int fg d\mu &= \int \left(\text{sign}g|g|^{q-1}\right) g d\mu \\ = \int \left(\text{sign}g \cdot g\right)|g|^{q-1} d\mu &= \int |g||g|^{q-1} d\mu \\ &= \int |g|^q d\mu = ||g||_q^q = 1 \end{aligned} \tag{52}$$

Assume then for the other direction that $\int fg d\mu = 1$, we will show that $f(x) = \text{sign}|g(x)|^{q-1}$, where equality here is equality almost everywhere.

By applying Hölder's inequalty as well as monotonicity, we get the following:

$$1 = \int fg d\mu \leq \int |fg| d\mu = ||fg||_1 \leq ||f||_p ||g||_q = 1 \tag{53}$$

Thus $||fg|| = ||f||_p ||g||_q$. Using **Theorem 6.2** from Folland then, we know that equality holds in Hölder's inequality if and only $|f|^p = C|g|^q$ almost everywhere for some constant $C > 0$.

Thus we take $|f|^p = C|g|^q$ almost everywhere for some $C > 0$. Note that as p, q dual we have $\frac{1}{p} + \frac{1}{q} = 1 \to \frac{q}{p} = q - 1$. We can therefore reorganize to get the following:

$$|f| = (C|g|^q)^{\frac{1}{p}} = C^{\frac{1}{p}}|g|^{\frac{q}{p}} = C^{\frac{1}{p}}|g|^{q-1} \tag{54}$$

Recall then that we showed in (48) that $\int fg d\mu = \int |fg| d\mu$. It is not particularly difficult to show that this implies $fg = |fg|$ a.e., which can seen in the following:

$$\begin{aligned} \int fg d\mu &= \int |fg| d\mu \\ \to \int (fg)^+ - (fg)^- d\mu &= \int (fg)^+ + (fg)^- \\ \to \int (fg)^+ d\mu - \int (fg)^- d\mu &= \int (fg)^+ d\mu + \int (fg)^- d\mu \\ &\to \int (fg)^- = 0 \end{aligned} \tag{55}$$

Which clearly implies $(fg)^- = 0$ a.e., i.e. fg is at most negative on a set of measure zero, so clearly $fg = |fg|$ almost everywhere. With this in mind, note we additionally have following:

$$1 = ||f||_p^p = \int |f|^p = \int C|g|^q = C \int |g|^q = C||g||_q^q = C \tag{56}$$

So $C = 1$, i.e. $|f| = |g|^{q-1}$ almost everywhere. With this, we can finally put everything together, also using the fact that $g = \text{sign}g|g|$:

$$\begin{aligned} |f| &= |g|^{q-1} \text{ a.e.} && (57)\\ \to |fg| = |g|^q \text{ a.e.} \quad &(\text{Multiplying both sides by } |g|^q) && (58)\\ \to fg = |g|^q \text{ a.e.} \quad &(\text{Using that } fg = |fg| \text{ a.e.}) && (59)\\ \to f = \frac{|g|^q}{g} \text{ a.e.} \quad &(\text{Dividing by } g) && (60)\\ & && (61) \end{aligned}$$

Note then that as $\frac{1}{\text{sign}g} = \text{sign}g$, we have that:

$$\frac{1}{g} = \frac{1}{\text{sign}g|g|} = \frac{\text{sign}g}{|g|} \tag{62}$$

Using this with (55) finally gets:

$$f = \frac{|g|^q}{g} = \frac{|g|^q \text{sign}g}{|g|} = \text{sign}g|g|^{q-1} \text{ a.e.} \tag{63}$$

Note here some of our work relies on $g \neq 0$ so that division is defined, but we note when $g(x) = 0$ we have $f(x) = 0$, which is still consistent with $f(x) = \text{sign}g(x)|g(x)|^{q-1}$. □

(b) Suppose that $||f_{t'}||_p = 1$ for some $0 < t' < 1$. Find some $g \in L^q$, $||g||_q = 1$, so that:

$$\int f_{t'} g d\mu = 1 \tag{64}$$

and let $F(t) = \int f_t g d\mu$. Observe as a result that $F(t) = 1$ for all $0 \leq t \leq 1$. Conclude that $f_t = f_0$ for all $0 \leq t \leq 1$.

Proof. Let $g = \text{sign}f_{t'}|f_{t'}|^{p-1}$. We note here that this has $|g| = |f_{t'}|^{p-1}$, where we get the following:

$$\int |g|^q d\mu = \int |f_{t'}|^{qp-q} \tag{65}$$

As p, q dual, we note that we have $q + p = qp$. Thus $qp - q = p$, and so we continue further:

$$\int |g|^q d\mu = \int |f_{t'}|^{qp-q} d\mu = \int |f_{t'}|^p = ||f_{t'}||_p^p = 1 \tag{66}$$

Thus $||g||_q^q = 1$, and so $||g||_q = 1$. Of course this also has $g \in L^q$, so this is the desired g.

We turn now to the given function $F(t)$, which we can examine in the following way:

$$\begin{aligned} F(t) &= \int f_t g d\mu = \int (1-t) f_0 g + f_1 g d\mu \\ &= (1-t) \int f_0 g d\mu + t \int f_1 g d\mu \end{aligned} \tag{67}$$

But $\int f_0 g d\mu$ and $\int f_1 g d\mu$ are constants, so it follows $F(t)$ is a linear polynomial in t. Moreover, we observe through an application of Hölder's inequality (and the triangle inequality) that we have the following:

$$\begin{aligned} F(t) = \int f_t g d\mu \leq \int |f_t g| d\mu = ||f_t g||_1 \leq ||f_t||_p ||g||_q \\ = ||f_t||_p = ||(1-t)f_0 + tf_1||_p \leq (1-t)||f_0||_p + ||f_1||_p = 1 - t + t = 1 \end{aligned} \tag{68}$$

Putting it all together, we have that $F(t)$ is a linear polynomial in t, $F(t) \leq 1$ on $[0,1]$, and $F(t') = 1$ for $t' \in (0,1)$.

As $F(t)$ is a linear polynomial, it is either strictly increasing/decreasing or constant, as observed by taking its derivative. However, we note either of the strictly increasing/decreasing cases are not possible.

In particular, this would have that for some small $\epsilon > 0$, $F(t' - \epsilon)$ or $F(t' + \epsilon)$ would be greater than $F(t')$ (depending on decreasing or increasing), but this would contradict that $F(t) \leq 1 = F(t')$ on $[0,1]$.

Thus the only possibility is that $F(t)$ is a constant function equal to 1.

We conclude that $f_t = f_0$ for all $0 \leq t \leq 1$ using part (a) then; in particular, note that having $F(t) = 1$ gives us that $||f_t||_p = 1$ for all $t \in [0,1]$ using our argument in (68).

Thus all the hypotheses of (a) are satisfied, so it is that $f_t = \text{sign} g |g|^{q-1} = f_0$, applying (a) to an arbitrary t and 0. This gives the desired conclusion. ☐

(c) Show that the strict convexity fails when $p = 1$ or $p = \infty$. What can be said about these cases?

Proof. We present two counterexamples for the cases where $p = 1, p = \infty$. In the case of $p = 1$, we consider on $L^1([-1,1])$ the following functions:

$$f_0 = \frac{1}{2}, f_1 = |x|$$

Clearly we have $||f_0||_1 = ||f_1|| = 1$. However, we consider the following for some $0 < t < 1$:

$$\begin{aligned} ||f_t||_1 = \int_{-1}^{1} |f_t| d\mu = \int_{-1}^{1} f_t d\mu \\ = \int_{-1}^{1} (1-t)\frac{1}{2} + t|x| d\mu = (1-t) \int_{-1}^{1} \frac{1}{2} + t \int_{-1}^{1} |x| \\ (1-t) + t = 1 \end{aligned} \tag{69}$$

And so strict convexity is violated.

For the case $p = \infty$, we work on $L^\infty([-1,1])$. We consider the following functions:

$$f_0 = \mathbb{1}_{[-1,0]}, f_1 = 1 \tag{70}$$

We clearly have $||f_0||_\infty = ||f_1||_\infty = 1$. However, considering some point $x \in [-1,0]$, we note for some $0 < t < 1$:

$$f_t(x) = (1-t)\mathbb{1}_{[-1,0]}(x) + t = (1-t) + t = 1 \tag{71}$$

We know $[-1,0]$ is positive measure then, so it follows $||f_t||_\infty \geq 1$, but this also violates strict convexity.

As for these cases, all we can say is that we have convexity, which follows from Minkowski's inequality. $\square$