Homework #1

Eric Tao
Math 237: Homework #1

February 10, 2024

Question 2. Let E be a normed vector space over $\mathbb{R}$. We call a subspace $H \subseteq E$ a hyperplane if the quotient space E/H has dimension 1.

2.1) Show that the closure of any subspace of E is also a subspace of E. Conclude that a hyperplane H is either closed or dense in E.

2.2) Let u be a linear functional on E. Prove that u is discontinuous if and only if there exists a sequence $\{x_n\}$ in E that converges to 0 such that $u(x_n) = 1$ for all n.

2.3) Let $x_0 \in E$ be a unit norm vector, and define H as the complement of the span of x_0. Show that every $x \in E$ can be uniquely decomposed as $x = t(x)x_0 + y(x)$ where $t : E \to \mathbb{R}$, and $y : E \to H$, linear. Further, prove that t, y are continuous if and only if H is closed.

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

Solution. 2.1)

Let $S \subset E$ be a vector subspace, and denote $\overline{S}$ as its closure. Of course, if S is closed, then $\overline{S} = S$, and therefore, the closure is a vector space.

Now, suppose $S \neq \overline{S}$. Then, we may describe $\overline{S}$ as the union of S and the limit points of S in E. Since $0 \in S \subset \overline{S}$, we need only show that $\overline{S}$ is closed under addition and scalar multiplication.

To check addition, we may discard the case where $x, y \in S$, as S is already a vector space. Thus, suppose x is a limit point of S, and $y \in S$. Since x is a limit point, there exists a sequence $\{x_n\} \subset S$ such that $\lim_{n\to\infty} x_n = x$. Then, consider the sequence $\{x_n + y\}$. Clearly, since $x_n \to x$, we have that $\lim_{n\to\infty} x_n + y = x + y$. Since $x \notin S$, being a limit point, $x + y$ cannot be in S, and hence, is a limit point of S. Hence, $x + y \in \overline{S}$. Without too much trouble, we see that the same argument holds when y is a limit point, where we leverage the sequences $\{x_n\}, \{y_n\}$ and consider their sum $\{x_n + y_n\}$.

Similarly, we can just check $x \notin S$ for scalar multiplication; if x is a limit point, $\{x_n\} \to x$, then of course $\{ax_n\} \to ax$ for $a \in \mathbb{R}$, and therefore, if $x \in S$, $ax \in S$. Thus, we have that $\overline{S}$ is closed under addition and scalar multiplication, and contains 0. Therefore, $\overline{S}$ is a vector subspace of E.

Now, let H be an arbitrary hyperplane. Of course, if H is closed, $\overline{H} = H$. So suppose H is not closed, and therefore $H \subset \overline{H}$ is a proper subset. Looking at E/H, since this has dimension 1, fixing some $z \in E \setminus H$, we may identify E/H as the span of $z + H$. Since $\overline{H}$ is a proper superset of H, there exists a $z' \in \overline{H}$ that does not belong to H. Under the projection into E/H, $\pi(z') = \alpha z + H$ for some $\alpha \in \mathbb{R} \setminus 0$, as otherwise, $z' \in H$, hence there exists a $h \in H$ such that $\alpha z + h = z'$ in E. Rearranging, this implies that $z = \frac{1}{\alpha}(z' - h)$. But, since $\alpha \in \mathbb{R}$, $z', h \in \overline{H}$, this implies that $z \in \overline{H}$. Hence, we have that $\overline{H} = E$. Since the closure of H in E is E, we have that H is dense in E, and we are done.

2.2)

First, we prove the forward direction. Suppose u is discontinuous. In particular then, it is discontinuous at the identity, since u is continuous if and only if it is continuous at the origin. Then, there exists some fixed $\epsilon > 0$, such that we may find a x_n with that $\|x_n - 0\| < 1/n$ and with $|u(x_n) - u(0)| = u(x_n) > \epsilon$. Now, consider the modified sequence $\{\frac{x_n}{u(x_n)}\}$. We notice that since $u(x_n) > \epsilon$, that term by term, this sequence is smaller in norm than $\{\frac{x_n}{\epsilon}\}$. Furthermore, since $x_n \to 0$, $\frac{x_n}{\epsilon} \to 0$, since $\|\frac{x_n}{\epsilon}\| = \frac{1}{\epsilon}\|x_n\| < \frac{1}{\epsilon}\frac{1}{n}$, which goes to

0 as $n \to \infty$ for a fixed ϵ. Thus, $\frac{x_n}{\epsilon} \to 0$ and therefore, $\{\frac{x_n}{u(x_n)}\} \to 0$. On the other hand though, since u is linear, $u\left(\frac{x_n}{u(x_n)}\right) = \frac{1}{u(x_n)}u(x_n) = 1$, as desired.

On the other hand, the backwards direction follows fairly easily. Since we have a sequence $\{x_n\} \to 0$ with $u(x_n) = 1$ for all n, of course, u is discontinuous at 0, because for $\epsilon = 1/2$, for any $\delta > 0$, we can find an x_n such that $\|x_n\| < \delta$, but by definition, $u(x_n) = 1 > \epsilon$. Hence, u is discontinuous at some point, and thus discontinuous.

2.3)

By the description of H, we can identify E/H as spanned by x_0. Then, for any $x \in E$, we can consider its image under the projection $\pi : E \to E/H$, $\pi(x) = t(x)x_0 + H$, for some map $t : E \to \mathbb{R}$; moreover, since π is linear, so must be t. Then, we may identify $y(x) = x - t(x)x_0$. We notice that $\pi(y(x)) = \pi(x - t(x)x_0) = \pi(x) - t(x)\pi(x_0) = t(x)x_0 + H - t(x)x_0 + H = 0 + H$, hence $y(x) \in H$.

We see this decomposition as unique, as x maps to exactly one coset of E/H due to the injectivity of left addition, so t is distinct. The uniqueness of y follows from the uniqueness of t. We also notice in what follows, that t, y are either both continuous or both discontinuous due to the definition of y.

Now, suppose t, y are continuous. Then, we can identify H as the inverse image $t^{-1}(0)$. Since t is continous, $t^{-1}(0)$ is closed, hence $H = t^{-1}(0)$ is closed.

On the other hand, suppose t, y discontinuous. Then, by 2.2, there exists a sequence $\{x_n\} \subset E$ such tthat $t(x_n) = 1$, and $x_n \to 0$. By the previous work, we can reexpress this sequence via our decomposition as:

$$x_n = t(x_n)x_0 + y(x_n) = x_0 + y(x_n)$$

But, since $x_n \to 0$, this implies that $y(x_n) \to -x_0$. Then, $-x_0 \in \overline{H}$, and hence from the work in 2.1, since $\overline{H}$ is a vector subspace, H is dense, i.e. not closed. Therefore, by the contrapositive, H being closed implies that t and thus y is continuous.

2.4)

Let u be a linear functional on E.

If u is trivial, then the result is trivial, as then the kernel of u is E, always closed, and the trivial map is continuous, because then the preimage of 0 is all of E.

Now, suppose u is not trivial. Then, because the kernel has codimension 1, looking at E/H, we may find a representative $z + H$ such that E/H is the span of $z + H$. Then, via 2.3, we may decompose any $x \in E$ as $x = t(x)z + y(x)$.

Then, u acting on any x has the action of $u(x) = u(t(x)z + y(x)) = t(x)u(z)$. Since $u(z)$ is a constant, the continuity of $u(x)$ is equivalent to the continuity of t. But, by 2.3, the continuity of t is equivalent to the closure of H. Thus, we have that:

$$u \text{ continuous } \iff t \text{ continuous } \iff H \text{ closed}$$

exactly our desired result. □

Question 5. Let E be a Banach space.

5.1) Suppose $T \in L(E, E)$, with $\|I - T\| < 1$. Prove that T is invertible, and that the series $\sum_{n=0}^{\infty}(I - T)^n$ converges in $L(E, E)$ to T^{-1}.

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

Solution. 5.1)

Firstly, we use the fact that since E is complete, so is $L(E, E)$ from Folland 5.4. We notice, that by the definition of the norm, that $\sup\{\|(I - T)x\| : \|x\| = 1\} < 1$; denote it as c. Considering $(I - T)(I - T)(x)$, for $\|x\| = 1$, call $(I - T)x = y$. Clearly, $\|y\| \leq c$. Looking at $(I - T)(y) = \|y\|(I - T)\left(\frac{y}{\|y\|}\right)$, due to the operator norm again, we see that $\|(I-T)(\frac{y}{\|y\|}\| \leq c$. Hence, for all $\|x\| = 1$, we have that $\|(I-T)^2(x)\| \leq c^2$. Then, $\sup\{\|(I-T)(I-T)x\| : \|x\| = 1\} \leq c^2$. Proceeding inductively, by considering $(I-T)^n(x) = (I-T)(I-T)^{n-1}(x)$, and using the same argument on $(I-T)^{n-1}(x)$ as having norm at most c^{n-1} in the same way, we see that $\|(I-T)^n\| \leq c^n$.

Now, we consider the sum $\sum_{n=0}^{\infty} \|(I-T)^n\|$. By the observations above, we have that $\|(I-T)^n\| \leq \|I-T\|^n$. So, we have a sum:

$$\sum_{n=0}^{\infty} \|(I-T)^n\| \leq \sum_{n=0}^{\infty} \|(I-T)\|^n = \sum_{n=0}^{\infty} c^n = \frac{1}{1-c}$$

where we've additionally used the fact that $\|I\| = 1$, which is clear, and identified this as an infinite geometric series with ratio less than 1. Then, since this is an absolutely convergent sum, and $L(E,E)$ is complete, $\sum_{n=0}^{\infty}(I-T)^n$ converges.

Now, we wish to show that $T\sum_{n=0}^{\infty}(I-T)^n$ acts as the identity, where we note that because T commutes with its powers, and T commutes with I, that we can write it on the left or right without ambiguity.

First, we look at the partial sums. We claim that $\sum_{n=0}^{k} T(I-T)^n = -(I-T)^{k+1} + I$.

The base case is easy. For $k = 1$, we see that this sum is exactly:

$$TI + T(I-T) = T + T - T^2 = 2T - T^2 = -(I-T)^2 + I$$

Now, suppose this is true for up to $k = m$. Then, we have that:

$$\sum_{n=0}^{m+1} T(I-T)^n = \sum_{n=0}^{m} T(I-T)^n + T(I-T)^{m+1} = -(I-T)^{m+1} + I + T(I-T)^{m+1} = (I-T)^{m+1}(-I+T) + I = -(I-T)^{m+2} + I$$

as desired. Then, to compute $T\sum_{n=0}^{\infty}(I-T)^n$, we can take the following limit:

$$\lim_{m\to\infty} T\sum_{n=0}^{m}(I-T)^n = \lim_{m\to\infty} -(I-T)^{m+2} + I$$

and because of the the work done with the norm, since $\|-(I-T)^{m+2}\| \leq \|I-T\|^{m+2}$, this goes to the 0 map as $m \to \infty$. Hence:

$$\lim_{m\to\infty} T\sum_{n=0}^{m}(I-T)^n = I$$

and hence, T is bijective with $\sum_{n=0}^{\infty}(I-T)^n$ as a left and right inverse, with the sum bounded.

5.2)

We consider the related operator $T^{-1}(S-T) = T^{-1}S - I$. By adapting the argument in the first part of 5.1, we see that $\|T^{-1}(S-T)\| \leq \|T^{-1}\|\|S-T\|$, where we do the same trick on considering $T^{-1}[(S-T)(x)]/\|(S-T)(x)\|$. So, we have that:

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

Thus, by 5.1 then, $T^{-1}S$ is invertible. But T is already invertible, and the composition of invertible bounded linear operators is invertible (as composition of bijective is bijective, composition of bounded is still bounded pretty easily: $\|f \circ g(x)\| \leq c_f\|g(x)\| \leq c_f c_g\|x\|$, and invertibility comes from, for $f \circ g$, considering $g^{-1} \circ f^{-1}$). Hence, $T \circ T^{-1}S = S$ is invertible.

Thus, we have shown that there exists open ball around any invertible operator T in $B(E,E)$ composed of invertible operators. Hence, by the local criterion for an open set, the set of invertible operators in $B(E,E)$ is open.

$\square$

Question 8. Suppose that $\mathcal{H}$ is a Hilbert space, $T \in L(\mathcal{H}, \mathcal{H})$.

8.1) Show that there exists a unique element that we denote $T^* \in L(\mathcal{H}, \mathcal{H})$ such that $\langle Tx, y\rangle = \langle x, T^* y\rangle$ for all $x, y \in \mathcal{H}$. Call T^* the adjoint of T.

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

8.3) Prove that $\|T^*\| = \|T\|$, $\|TT^*\| = \|T\|^2$, $(aS + bT)^* = \overline{a}S^* + \overline{b}T^*$, $(ST)^* = T^*S^*$, and $T^{**} = T$.

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

8.5) Show that T is unitary if and only if T is invertible, with $T^{-1} = T^*$.

Solution. 8.1)

Suppose there exists another $T' \in L(\mathcal{H}, \mathcal{H})$ such that $\langle x, T'y\rangle = \langle Tx, y\rangle = \langle x, T^*y\rangle$.

Then, we consider $\langle x, T'y\rangle - \langle x, T^*y\rangle = 0$. By conjugate symmetry, we have that:

$$\overline{\langle T'y, x\rangle} - \overline{\langle T^*y, x\rangle} = 0$$

But, the complex conjugate distributes over addition, so:

$$\overline{\langle T'y, x\rangle - \langle T^*y, x\rangle} = 0$$

Now, using linearity of the first term, we have that:

$$\overline{\langle T'y - T^*y, x\rangle} = 0$$

Since x is arbitrary, we may choose x as $T'y - T^*y$. Since the inner product is real in this case, and as a Hilbert space, extends to a norm, we have that

$$\langle T'y - T^*y, T'y - T^*y\rangle = 0 \implies \|T'y - T^*y\|^2 = 0 \implies \|T'y - T^*y\| = 0$$

Hence, by the properties of the norm, we have that $T'y - T^*y = 0 \implies T^*y = T'y$. Since y was arbitrary, this implies that $T' = T^*$ on all of $\mathcal{H}$.

8.2)

We consider the action of $V^{-1}T^\dagger V$ on a test vector y. By definition, $V(y) = f_y \in \mathcal{H}^*$, which acts via $f_y(x) = \langle x, y\rangle$. Then, again by definition, $T^\dagger$ acts on $f_y(x)$, sending it to the functional that acts via $\tilde{f}_y(x) = \langle T(x), y\rangle$. Lastly, V^{-1} takes $\tilde{f}_y$ and sends it back to $\mathcal{H}$ to z, such that z is the unique element in $\mathcal{H}$ such that $\langle x, z\rangle = \langle T(x), y\rangle$, due to the definition of $\tilde{f}_y$. But, letting x, y range over $\mathcal{H}$, this is exactly the action of T^*. Since T^* is unique, this is an equality of operators.

8.3)

First, we prove that $(T^*)^* = T$. Let x, y be arbitrary elements of $\mathcal{H}$, and consider the equation $\langle T^*x, y\rangle = \langle x, (T^*)^*(y)\rangle$. We have that following string of equalities:

$$\overline{\langle Ty, x\rangle} = \overline{\langle y, T^*x\rangle} = \langle T^*x, y\rangle = \langle x, (T^*)^*y\rangle = \overline{\langle (T^*)^*y, x\rangle}$$

which implies then that $\langle Ty, x\rangle = \langle (T^*)^*y, x\rangle \implies \langle [T - (T^*)^*](y), x\rangle = 0$ for all x, y. Then, yet again, with the same trick of choosing $x = [T - (T^*)^*](y)$, we see that $T - (T^*)^* = 0$ as operators, and thus $T = (T^*)^*$.

Next, we prove a statement on $V : \mathcal{H} \to \mathcal{H}^*$ that sends $x \mapsto f_x(y) = \langle y, x\rangle$. First, let y be any unit norm vector, and we will consider the norm of f_y. Let x be yet another unit norm vector. Then, by the Cauchy-Schwarz inequality, we have that:

$$\|f_y(x)\| = |\langle x, y\rangle| \leq \|x\|\|y\| \leq 1$$

where we have used the fact that $\|x\|, \|y\| = 1$. Furthermore, by choosing $x = y$, we see that this attains 1. Thus, we have that $\|f_y\| = 1$. Since this is true for all y, we may conclude that $\|V\| = 1$. Considering the fact that $V^{-1} \circ V$ acts on identity on $\mathcal{H}$ (or, equivalently, $V \circ V^{-1}$ on $\mathcal{H}$), we can conclude that $\|V^{-1}\| = 1$.

Finally, we look at $\|T^\dagger\|$. Letting f be a unit norm vector in $\mathcal{H}^*$. Via the isomorphism that identifies $y \in \mathcal{H}$ with $f_y(x) = \langle x, y\rangle$, it is clear that $\|y\| = 1 \iff \|f_y\| = 1$ due to Cauchy-Schwarz. Suppose $\|f_y\| = 1$. Then, for $x \in \mathcal{H}$ with unit norm, we have that:

$$|f_y(x)| \leq \|x\|\|y\| \leq \left\|\frac{y}{\|y\|}\right\| \|y\| = \|y\|$$

where Cauchy-Schwarz guarantees that we achieve equality at $\frac{y}{/}\|y\|$. Then, since this inequality holds for all x, and is independent of x, we see that $\|f_y\| = \|y\|$.

In any case, looking at the action of $T^\dagger$ on f_y, let x be a unit norm vector in $\mathcal{H}$, then we see that

$$\|T^\dagger f_y(x)\| = \|f_y(T(x))\| = |\langle T(x), y\rangle| \leq \|T(x)\|\|y\| \leq \|T\|$$

where we use the fact that x, y have unit norm. Thus, we may conclude that $\|T^\dagger\| \leq \|T\|$.

Then, using the same argument as used in 5.1 for showing that the operator norm is submultiplicative, we see that:

$$\|T^*\| = \|V^{-1}T^\dagger V\| \leq \|V^{-1}\|\|T^\dagger\|\|V\| = \|T^\dagger\| \leq \|T\|$$

However, we already have that $T = (T^*)^*$, so we may run this same argument with $\|T\| = \|(T^*)^*\| = \|(V')^{-1}(T^*)^\dagger V'\| \leq \|T^*\|$ with V' as the isomorphism from $\mathcal{H}^* \to \mathcal{H}$ in the same way. Thus, we have that $\|T\| = \|T^*\|$.

Now, let x have unit norm. Then, we look at the following string of inequalities:

$$\|Tx\|^2 = \langle Tx, Tx\rangle = \langle x, T^*Tx\rangle \leq \|x\|\|T^*Tx\| \leq \|x\|\|T^*T\|\|x\| = \|T^*T\|$$

where we notice since $\langle Tx, Tx\rangle = \langle x, T^*Tx\rangle$, the right side is positive and real, and thus is equal to its absolute value, where we use Cauchy-Schwarz.

Since this is true for all x with unit norm, this implies $\|T\|^2 \leq \|T^*T\|$. But by submultiplicativity, we have that $\|T^*T\| \leq \|T^*\|\|T\| = \|T\|^2$ from $\|T^*\| = \|T\|$. Hence, $\|T^*T\| = \|T\|^2$. We will see later that since $(T^*T)^* = T^*T$, and $\|T\| = \|T^*\|$ will show this to be equivalent to the problem statement.

To see $(aS + bT)^* = \overline{a}S^* + \overline{b}T^*$ is easy via the conjugate linearity of V, as clearly, V^{-1} must be conjugate linear itself since if we consider $kf_y(x) = k\langle x, y\rangle = \langle x, \overline{k}y\rangle$, evidently, $V(\overline{k}y) = kf_y$, and so $V^{-1}(kf_y) = \overline{k}f_y$. We see that:

$$(aS + bT)^*(y) = V^{-1}(aS + bT)^\dagger V(y) = V^{-1}(aS + bT)^\dagger f_y = V^{-1}(f_y \circ (aS + bT)) =$$
$$V^{-1}[a(f_y \circ S) + b(f_y \circ T)] = \overline{a}V^{-1}f_y \circ S + \overline{b}V^{-1}f_y \circ T = \overline{a}S^* + \overline{b}T^*(y)$$

since this is true for arbitrary $y \in \mathcal{H}$, this is an equality of operators.

Similarly:

$$(ST)^*(y) = V^{-1}(ST)^\dagger V(y) = V^{-1}(ST)^\dagger f_y$$

Considering an arbitrary $x \in \mathcal{H}$, we see that:

$$(ST)^\dagger f_y(x) = f_y(ST(x)) = S^\dagger f_y(T(x)) = T^\dagger \circ S^\dagger \circ f_y(x)$$

Since this is true for all x, y, we have that:

$$(ST)^* = V^{-1} \circ T^\dagger \circ S^\dagger \circ V$$

On the other hand, by definition, we have that:

$$T^* S^* = (V^{-1} \circ T^\dagger \circ V) \circ (V^{-1} \circ S^\dagger \circ V) = V^{-1} \circ T^\dagger \circ I \circ S^\dagger \circ V = V^{-1} \circ T^\dagger \circ S^\dagger \circ V$$

completing our proof.

8.4)

Recall that the definition of $R(T)^\perp = \{x \in \mathcal{H} : \langle x, y\rangle = 0 \forall y \in R(T)\}$.

First, suppose $y \in R(T)^\perp$. Then, by definition, we have that $\langle y, Tx\rangle = \langle Tx, y\rangle = 0$ for all $x \in \mathcal{H}$. Then, we have that $\langle x, T^* y\rangle = 0$. Specifically, this must be true for $x = T^* y$, which implies that $T^* y = 0$. Thus, $R(T)^\perp \subseteq N(T^*)$.

Next, suppose $y \in N(T^*)$. Then, we have that $\langle x, T^*(y)\rangle = 0$ for all x, which we can see by the Schwarz inequality, and how $\|T^* y\| = 0$. Then, we have that $\langle T(x), y\rangle = 0$ for all $x \in \mathcal{H}$, which implies that $\langle y, T(x)\rangle = 0$, and thus by definition again, $y \in R(T)^\perp$.

Now, from the first part, we have that:

$$N(T)^\perp = N(T^{**})^\perp = (R(T^*)^\perp)^\perp$$

It should be clear that for X a subset, that $X \subset (X^\perp)^\perp$, as for any $x \in X$, we have that:

$$\langle y, x\rangle = 0 = \overline{\langle x, y\rangle} = \langle x, y\rangle$$

for any $y \in X^\perp$. However, we see that the last expression is exactly the defining statement of $(X^\perp)^\perp$. Hence, $X \subset (X^\perp)^\perp$. So, we have that $R(T^*) \subseteq N(T)^\perp$. In particular, from problem 56, this implies that $N(T)^\perp$ is the smallest closed subspace that contains $R(T^*)$. But from problem 2.1, since $R(T^*)$ is a subspace, $\overline{R(T^*)}$ is a subspace, hence the smallest closed subspace, hence equal to $N(T)^\perp$.

Folland #56:

Let E be a subset of $\mathcal{H}$. Then $(E^\perp)^\perp$ is the smallest closed subspace containing E.

We have already shown that $E \subset (E^\perp)^\perp$. From Proposition 5.21 in Folland, we know that any subset $E^\perp$ is closed. Moreover, from the linearity of the inner product in the first argument, of course this is a vector subspace of $\mathcal{H}$. Thus, we need only prove that it is the smallest such closed subspace.

Suppose we have another closed subspace of $\mathcal{H}$, call it F such that $E \subseteq F$. Then, of course, $F^\perp \subseteq E^\perp$, since if we're orthogonal to all of F, and F contains E, then we're orthogonal to E. Evidently then, $(E^\perp)^\perp \subseteq (F^\perp)^\perp$, substituting $E^\perp$ for F, and $F^\perp$ for E above.

Suppose we fix some element $x \in (F^\perp)^\perp$. By theorem 5.24 in Folland, since F is a closed subspace, then we can rewrite $\mathcal{H} = F \oplus F^\perp$, and hence, $x = f + f'$ for $f \in F, f' \in F^\perp$. But, of course, $0 = \langle x, f'\rangle = \langle f + f', f'\rangle = \langle f, f'\rangle + \langle f', f'\rangle = \langle f', f'\rangle$, which implies that $f' = 0$. Hence, $x = f$. Since we can do this for all $x \in (F^\perp)^\perp$, this implies that $(F^\perp)^\perp \subseteq F$. Hence, $(E^\perp)^\perp \subseteq (F^\perp)^\perp \subseteq F$, and therefore, must be the smallest such closed subspace.

8.5)

The backward direction is easy. We have that:

$$\langle x, y\rangle = \langle T^{-1} T x, y\rangle = \langle T^* T x, y\rangle = \langle Tx, T^{**} y\rangle = \langle Tx, Ty\rangle$$

for all $x, y \in \mathcal{H}$.

On the other hand, suppose T is unitary. Then, we have that:

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

Since T is invertible, we can in particular, choose $x = T^{-1}(z)$. Then, we have that:

$$\langle T^{-1}z, y\rangle = \langle T^*TT^{-1}(z), y\rangle = \langle T^*z, y\rangle$$

Since z ranges over all of $\mathcal{H}$ as T is invertible, we can conclude that $T^{-1} - T^* = 0 \implies T^{-1} = T^*$ everywhere.

□

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

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

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

12.3) Show that the dimension of M is at most C^2, by proving that if $\{f_k\}$ is any orthogonal sequence in M, then $\sum_k |f_k(x)|^2 \leq C^2$ for all $x \in [0,1]$.

Solution. 12.1)

Consider the inclusion as vector spaces $i : M \to C([0,1])$. Evidently, this map is linear, as the addition and scalar multiplication in L^2 and $C([0,1])$ act in the same way. Then, we wish to show it as closed.

Let $\{f_n\} \to f$ be a convergent sequence of functions in M, such that $\{i(f_n)\} \to y \in C([0,1])$. Suppose $i(f) = f \neq g$. Then, there must exist some x_0 such that $|f(x_0) - g(x_0)| > 0$. By continuity then, since $f - g$ is continuous as well, there exists a $\epsilon > 0$ such that for all $|x - x_0| < \delta$, $|f(x) - g(x)| > \epsilon$. Note that in the case $x_0 - \delta < 0$ or $x_0 + \delta > 1$, we adjust δ to be the smaller of δ and the distance to the endpoint. Thus, we have then that:

$$\|f - g\|_2 = \sqrt{\int_{[0,1]} |f(x) - g(x)|^2 dx} \geq \sqrt{\int_{[x_0-\delta, x_0+\delta]} |f(x) - g(x)|^2 dx} \geq \sqrt{2\delta\epsilon^2}$$

On the other hand, we have that:

$$\|f - g\|_2 = \|f - f_n\|_2 + \|f_n - g\|_2$$

Since $f \to f_n$ in the L^2 norm, we may choose N_1 such that for all $n > N_1$, $\|f - f_n\|_2 < \epsilon\sqrt{2\delta}/2$.

Looking at $\|f_n - g\|_2 = \sqrt{\int_{[0,1]} |f_n - g|^2} \leq \sqrt{\int_{[0,1]} \|f_n - g\|_u^2}$, since $f_n \to g$ in the uniform norm, we may choose N_2 such that for all $n > N_2$, $\|f_n - g\|_u < \epsilon\sqrt{2\delta}/2$.

Then, choosing $n > \max(N_1, N_2)$, we see that:

$$\|f - g\|_2 = \|f - f_n\|_2 + \|f_n - g\|_2 < \epsilon\sqrt{2\delta}/2 + \sqrt{\left(\epsilon\sqrt{2\delta}/2\right)^2} = \epsilon\sqrt{2\delta}$$

Thus, $\epsilon\sqrt{2\delta} < \epsilon\sqrt{2\delta}$, a contradiction. Hence, $f = g$. Therefore, the inclusion is a closed map. Moreover, $C([0,1])$ is a Banach space under $\|\cdot\|_u$. Moreover, since M is a closed subspace of a Banach space, it is itself a Banach space with the same norm. Hence, by the closed graph theorem (5.12, Folland), we have that because the inclusion is a closed linear map, then it is bounded.

By the definition of a bounded linear map then, we have that there exists a $C > 0$ such that $\|i(f)\|_u = \|f\|_u \leq C\|f\|_2$.

12.2)

First, we note that since M is a closed subspace of a Hilbert space, it too is a Hilbert space with the same inner product as L^2, restricted to M.

Consider the map that takes a function in M and evaluates it at a point $x \in [0,1]$. Denote this map as $T_x : M \to F$, for F our base field.

Clearly, this map is linear, since $T_x(af+bg) = (af+bg)(x) = af(x)+bg(x)$, due to how addition and scalar multiplication of functions is defined pointwise. Moreover, of course, $|T_x(f)| = |f(x)| \leq |f|_u$, as the uniform norm is the supremum over all $x \in [0,1]$. But, by 12.1, this is at most $C\|f\|_2$. Hence, T_x is bounded. Since T_x is a bounded linear functional, it belongs to M^*. But then, by Theorem 5.25 (Folland), there exists a unique $g_x \in M$ such that $f(x) = T_x(f) = \langle f, g_x \rangle$, for all $f \in M$.

In particular, we have that:

$$\|g_x\|_2^2 = \langle g_x, g_x \rangle = T_x(g_x) = g_x(x) \leq \|g_x\|_u \leq C\|g_x\|_2$$

Assuming first that $\|g_x\|_2^2 \neq 0$, this implies after dividing both sides by $\|g_x\|_2$, that:

$$\|g_x\|_2 \leq C$$

and we notice that if $g_x = 0$, then this inequality is still satisfied.

12.3)

Let $\{f_k\}$ be an orthogonal sequence in M. We may replace this with an orthonormal sequence by replacing f_k with $f_k/\|f_k\|_2$. Further, restrict to a finite sequence, restricting to a subsequence if need be - say that $\{f_k\}_{k=1}^N$ is our orthonormal subsequence. Fix an $x \in [0,1]$. By 12.2, there exists g_x such that $f(x) = \langle f, g_x \rangle$ for all $f \in M$. Thus, we have that:

$$\sum_{k=1}^{N} |f_n(x)|^2 = \sum_{k}^{N} |\langle f_n, g_x \rangle|^2$$

Now, by Bessel's Inequality, after using the fact that $|\langle f_n, g_x \rangle|^2 = |\langle g_x, f_n \rangle|^2$, since the modulus of the transpose is equal to the original modulus:

$$\sum_{k=1}^{N} |\langle g_x, f_k \rangle|^2 \leq \|g_x\|_2^2$$

and by 12.2, we have that this quantity is at most C^2. Hence, we have that:

$$\sum_{k=1}^{N} |f_k(x)|^2 \leq C^2$$

for all $x \in [0,1]$.

Then, we have that:

$$\sum_{k=1}^{N} \|f_k\|_2^2 = \sum_{k=1}^{N} \int_{[0,1]} |f_k|^2 dx = \int_{[0,1]} \sum_{k=1}^{N} \sum_{k=1}^{N} |f_k|^2 \leq \int_{[0,1]} C^2 = C^2$$

On the other hand, using the normality, we have that:

$$\sum_{k=1}^{N} \|f_k\|_2^2 = N$$

Hence, we have that $N \leq C^2$, and any finite sequence of orthonormal vectors has at most C^2 vectors. Since any orthogonal sequence must give rise to a orthonormal sequence, by dividing out by a norm, and we may always look at finite subsequences of infinite sequence, this must be true for arbitrary orthogonal sequences of $\{f_k\}$. Thus, the maximal number of distinct elements in an orthogonal sequence is C^2, hence the dimensionality of M is at most C^2 as a vector space.

□

Question 20. Recall that L^p denotes the space of real-valued functions such that their p-th power is integrable. Suppose that $\|f_0\|_{L_p} = \|f_1\|_{L_p} = 1$. Define

$$f_t = (1-t)f_0 + tf_1$$

Of course, $\|f_t\|_{L_p} < 1$ for all $t \in (0,1)$ unless $f_0 = f_1$.

20.1)

Let $f \in L^p, g \in L^q$, with $1/p + 1/q = 1$, $\|f\|_{L^p} = 1, \|g\|_{L^q} = 1$. Show that if

$$\int fg d\mu = 1$$

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

20.2)

Suppose that $\|f_{t'}\|_{L^p} = 1$ for some $0 < t' < 1$. Find $g \in L^q$ with $\|g\|_{L^q} = 1$, such that:

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

and denote $F(t) = \int f_t g d\mu$. Prove that $F(t) = 1$ for all $t \in [0,1]$, and conclude that $f_t = f_0$ for all $t \in [0,1]$.

20.3)

Show that this fails when $p = 1, p = \infty$. What can we say in these cases?

Solution. 20.1)

Clearly, we have the following string of inequalities:

$$\int fg d\mu \leq \int |fg| d\mu \leq \|f\|_p \|g\|_q$$

where we identify $|fg| = \|fg\|_1$ when f, g are real-valued, and apply Hölder's inequality.

However, by hypothesis, $\int fg d\mu = 1$, and $\|f\|_p = 1 = \|g\|_q$. Hence, this implies that $\int |fg| d\mu = 1 = \|f\|_p \|g\|_q$.

Then, from 6.2 in Folland, we recall that equality in Hölder's inequality holds if and only if $|f|^p = c|g|^q$ almost everywhere for some non-0 constant c.

With some algebraic manipulation, we can see that since $1/p + 1/q = 1$, that $q/p + 1 = q \implies q/p = q - 1$.

$$|f| = (c|g|^q)^{1/p} = c^{1/p}|g|^{q/p} = c^{1/p}|g|^{q-1}$$

From the fact that $\int fg d\mu = 1 = \int |fg| d\mu$, we may conclude that $fg = |fg|$ a.e. Furthermore, we see that:

$$\|f\|_p^p = \int |f|^p d\mu = \int c|g|^q d\mu = c \int |g|^q d\mu = c\|g\|_q^q$$

which implies that $c = 1$.

Therefore, we can conclude that $f(x) = h(x)g(x)^{q-1}$ a.e., for $h(x) = \pm 1$. But again, since $fg = |fg|$, f, g must carry the same sign. Thus, we can safely replace $h(x) = \text{sign}(g(x)$, as desired.

20.2)

From 20.1), we take $g = \text{sign}(f_{t'}(x))|f_{t'}(x)|^{p-1}$. Evidently:

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

that is $\|g\|_q^q = 1 \implies \|g\|_q^q = 1$.

Furthermore, we may evaluate $\int f_{t'} g d\mu$:

$$\int f_{t'} g d\mu = \int f_{t'} \text{sign}(f_{t'}(x)) |f_{t'}(x)|^{p-1} d\mu = \int |f_{t'}|^p d\mu = \|f_{t'}\|_p^p = 1$$

where we use the fact that $f_{t'} * \text{sign}(f_{t'}) = |f_{t'}|$.

Now, looking at $F(t)$, we can expand to find that:

$$F(t) = \int f_t g d\mu = \int (1-t) f_0 g + t f_1 g d\mu = (1-t) \int f_0 g d\mu + t \int f_1 g d\mu$$

In particular, we see that with respect to t, $\int f_0 g d\mu$ and $\int f_1 g d\mu$ are constants, hence F is linear with respect to t.

Moreover, again from Hölder's inequality, we see that:

$$F(t) = \int f_t g d\mu \leq \int |f_t g| d\mu \leq \|f_t\|_p \|g\|_q = \|(1-t) f_0 + t f_1\|_p \leq (1-t)\|f_0\|_p + t\|f_1\|_p = 1$$

And, of course, at t', $F(t') = 1$. $F(t)$ then is a linear function that attains an extrema within the interior of its (connected) domain, $[0,1]$ and hence is constant; i.e. $F(t) = 1$.

Now, since $F(t) = 1$ for all $0 \leq t \leq 1$, we can conclude from 20.1 that $f_t = \text{sign}(g(x))|g(x)|^{q-1} = f_0$, for all t.

20.3)

Suppose $p = 1$, and take our set to be $[0,1]$. Then, we can consider the functions $f_0 = 1, f_1 = 2x$. Of course, these both have unit norm. Looking at the integral, we see that:

$$\|f_t\|_1 = \int |f_t| dx = \int (1-t) + 2tx dx = tx^2 + (1-t)x|_0^1 = t + (1-t) = 1$$

and hence, strict convexity fails.

Similarly, take $p = \infty$, and take our set to be any measurable set E. Let $A \subset E$ be a proper measurable subset, and take $f_0 = \chi_A, f_1 = 1$, where χ_A takes on 1 on A and 0 on $E \setminus A$. Again, $\|f_0\|_\infty = 1 = \|f_1\|_\infty$, clearly. Looking at any points $x \in A$, clearly, for any t, $|(1-t)f_0(x) + t f_1(x)| = |(1-t) + t| = 1$, and hence $\|f_t\|_\infty = 1$ for all $t \in (0,1)$, and strict convexity fails.

Thus, in such a case, we can only expect to have convexity, where convexity still follows from Minkowski's inequality.

$\square$