Functional Analysis F3/F4/NVP (2005)
Homework assignment 2

All students should solve the following problems:

1. Section 2.7: Problem 8.

2. Let $x_1(t) = t^2 e^{-t/2}$, $x_2(t) = te^{-t/2}$ and $x_3(t) = e^{-t/2}$. Orthonormalize x_1, x_2, x_3, in this order, in the Hilbert space $L^2[0, +\infty)$.

3. Section 3.9: Problem 6.

4. Section 4.3: Problem 14.

Students taking Functional Analysis as a 6 point course should also solve the following problems:

5. Let $T : H \to H$ be a linear operator on the Hilbert space H. Prove that T is unitary if and only if $T(M)$ is a total orthonormal set in H for each total orthonormal set M in H.

6. Let Y be a closed subspace of a normed space X and let $A : X'/Y^a \to Y'$ be the operator defined by $A(f + Y^a) = f_{|Y}$. Prove that A is an isomorphism of normed spaces.

[Notation: $f_{|Y}$ is the restriction of f to Y, see p.99 (middle). Y^a is the annihilator of Y as defined in Section 2.10, problem 13; furthermore, X'/Y^a is a normed space as in Section 2.3, problem 14.]

Solutions should be handed in by Wednesday, February 16, 18.00. (Either give the solutions to me directly or put them in my mailbox, third floor, House 3, Polacksbacken.)

Functional Analysis F3/F4/NVP

Comments to homework assignment 2

1. Note that from the definition (p. 91) of "bounded", our task in this problem is to show that there does *not* exist a constant c such that $||T^{-1}x|| \leq c||x||$ holds for all $x \in \mathcal{D}(T^{-1})$. Several students made the mistake of playing with some vector x for which $T^{-1}x$ is not defined, i.e. a vector outside the domain $\mathcal{D}(T^{-1}) = \mathcal{R}(T)$. For such vectors "$T^{-1}x$" is *nonsense*, and we cannot conclude anything about T^{-1} being bounded or not bounded by studying such vectors. Note also that for each *individual* vector $x \in \mathcal{D}(T^{-1})$ there *will* exist some constant c such that $||T^{-1}x|| \leq c||x||$ holds. Hence to prove that T^{-1} is not bounded, we have to make a clever choice of an *infinite sequence* of vectors $x_1, x_2, ...$ in $\mathcal{D}(T^{-1})$, and prove that there is no constant c which works for *all* these vectors. (There are also alternative approaches; but the point is that a *no* proof can work by studying just *one* explicit fixed vector x in $\mathcal{D}(T^{-1})$.)

4. Some students seem to have misunderstood the statement of Theorem 4.3-3: Note that the *function* $f : X \to K$ given by $f(x_0) = ||x_0||$ $(\forall x_0 \in X)$ is *not* a linear functional; for instance it does not satisfy $f(x+y) = f(x) + f(y)$ for all $x, y \in X$. What Theorem 4.3-3 says is that given some *fixed* vector $x_0 \in X$ $(x_0 \neq 0)$, there exists a bounded linear functional $f : X \to K$ such that $||f|| = 1$ and $f(x_0) = ||x_0||$. Note that this functional will *not* satisfy $f(x) = ||x||$ for *all* $x \in X$.

The same thing expressed with symbols: Theorem 4.3-3 says

$$\forall x_0 \in X - \{0\} : \quad \exists \tilde{f} \in X' : \quad ||\tilde{f}|| = 1, \quad \tilde{f}(x_0) = ||x_0||.$$

Theorem 4.3-3 does *not* say:

$$''\exists \tilde{f} \in X' : \quad \forall x_0 \in X - \{0\} : \quad ||\tilde{f}|| = 1, \quad \tilde{f}(x_0) = ||x_0||.''$$

Here is a speculation on what might be the origin of this misconception: In Theorem 4.3-3 it says "...and let $x_0 \neq 0$ be any element of X." This *should* be understood as "and let $x_0 \neq 0$ be a *fixed* (arbitrary) vector in X." However, some students might have read it as something like "...and we use the letter x_0 to denote an arbitrary (varying) non-zero element in X." With this interpretation it actually looks as if the theorem claims (falsely) that the *function* $f(x_0) = ||x_0||$ $(\forall x_0 \in X)$ is linear! However I think every professional mathematician would find the statement of the theorem perfectly clear and would say that the interpretation with "varying x_0" is *incorrect*. Ultimately it is a *convention*; mathematicians always read "let ... be any ..." as "let ... be a *fixed* ...". Unfortunately it seems that eg. professional physicists do *not* always use the same convention, so I definitely find the misconception understandable.

5. Some students have referred to Theorem 3.10-6(f) to conclude that T is unitary (after having proved that T is isometric and surjective. Unfortunately, Theorem 3.10-6(f) is only stated for *complex* Hilbert spaces; hence it is not strong enough for us (since we did not say anything about the ground field K in the problem formulation we wish the proof to hold both for $K = \mathbb{R}$ and for $K = \mathbb{C}$).

However, Theorem 3.10-6(f) *is* true also for *real* Hilbert spaces; this can be proved by mimicking the last part of our solution to problem 5 below (in particular, one uses polarization to show that $||Tx|| = ||x||, \forall x \in H$ implies $\langle Tx, Ty \rangle = \langle x, y \rangle, \forall x, y \in H$).

Here follows a discussion of some rather intricate issues in the problem. (The following matters did not affect the score by more than ± 1, and you may safely consider the following discussion as extracurricular!) The hardest part of the problem is to give the proof in the direction [$\forall M \subset H : M$ total orthonormal set $\Longrightarrow T(M)$ total orthonormal set] $\Longrightarrow$ [T unitary]. Note that in the statement of the problem *we did not assume that T is bounded*; this can be deduced from the assumption. (The boundedness is not an issue in the opposite direction since a unitary operator is bounded by definition.) Furthermore, note that even if M is a total orthonormal set and $T(M)$ is a total orthonormal set, we *may*, apriori, have $T(x) = T(y)$ for some vectors $x, y \in M$, $x \neq y$. (Eg. assume H is separable so that there is a total orthonormal sequence $e_1, e_2, ...$ in H; then $M = \{e_1, e_2, ...\}$ is a total orthonormal set. Assume $T(e_1) = e_1$, $T(e_2) = e_1$, $T(e_3) = e_2$, $T(e_4) = e_2$, $T(e_5) = e_3$, etc. Then $T(M) = \{e_1, e_2, ...\}$, a total orthonormal set!) This possibility can only be excluded by also considering some (appropriately chosen) *different* total orthonormal set M' in H (eg. if $T(x) = T(y)$ for $x, y \in M$, $x \neq y$, then apply the assumption to a total orthonormal set which contains the unit vector $2^{-\frac{1}{2}}(x - y)$).

6. As part of a completely correct solution you must prove that A is *well-defined*. In fact, this should be a reflex whenever a function on a set of equivalence classes is defined in a way using a choice of *representatives* for the equivalence classes! (In the present problem f is a representative for the equivalence class $f + Y^a$, and $A(f + Y^a)$ is defined *using* this representative f. Cf. also problem 5 in Homework no 1.) In the present problem it is fairly easy to see that A is indeed well-defined, but *this should be stated explicitly*.

Functional Analysis F3/F4/NVP
Solutions to homework assignment 2

1. As suggested in the hint we let T be the operator $T : \ell^\infty \to \ell^\infty$ defined by $T((\xi_1, \xi_2, \xi_3, ...)) = (\xi_1/1, \xi_2/2, \xi_3/3, ...)$. This operator is easily seen to be linear. We also have for all $(\xi_j) \in \ell^\infty$:

$$||T((\xi_j))|| = ||(\xi_1/1, \xi_2/2, \xi_3/3, ...)|| = \sup_j |\xi_j/j| \le \sup_j |\xi_j| = ||(\xi_j)||.$$

Hence T is bounded and $||T|| \le 1$. We note that the range of T is:

$$\begin{aligned}\mathcal{R}(T) &= \{T((\xi_j)) \mid (\xi_j) \in \ell^\infty\} \\ &= \{(\eta_1, \eta_2, \eta_3, ...) \mid \eta_j = \xi_j/j, \ \sup_j |\xi_j| < \infty\} \\ &= \{(\eta_1, \eta_2, \eta_3, ...) \mid \sup_j |j\eta_j| < \infty\}\end{aligned}$$

(Note: Up to here this work has been done in class, in a problem session where I solved problems 5,6 on p. 101.) Note that if $(\eta_j) = T((\xi_j))$ then we must have $\xi_j = j\eta_j$ for all j, i.e. $(\xi_j) = (j\eta_j)$. (This has already been used in the above computation of $\mathcal{R}(T)$.) This shows that T is injective, and that the inverse map $T^{-1} : \mathcal{R}(T) \to \ell^\infty$ is given by $T^{-1}((\eta_j)) = (j\eta_j)$. We will prove that T^{-1} is not bounded.

Given any $n \ge 1$, let e_n be the vector $e_n = (0, 0, ..., 0, 0, 1, 0, 0, ...)$, with the number 1 in position number n. We see from the above formula for $\mathcal{R}(T)$ that $e_n \in \mathcal{R}(T)$. Also $||e_n|| = 1$ and $T^{-1}(e_n) = (0, 0, ..., 0, 0, n, 0, 0, ...) = ne_n$, and hence $||T^{-1}(e_n)|| = ||ne_n|| = n$. Now *if* T^{-1} were bounded, then there would exist a constant $c \ge 0$ such that

$$(*) \qquad ||T^{-1}(x)|| \le c||x||, \qquad \forall x \in \mathcal{R}(T).$$

But then choose n as an integer *larger* than c; we then have $e_n \in \mathcal{R}(T)$ and $||T^{-1}(e_n)|| = n > c = c||e_n||$, which *contradicts* (*). Hence T^{-1} is *not* bounded.

Alternative solution. Some students have instead studied the operator $T : C[0, 1] \to C[0, 1]$ defined by $(Tx)(t) = \int_0^t x(s)\, ds$. This operator is linear, since for all $x_1, x_2 \in C[0, 1]$ and all $\alpha, \beta \in K$ we have

$$\begin{aligned}(T(\alpha x_1 + \beta x_2))(t) &= \int_0^t (\alpha x_1(s) + \beta x_2(s))\, ds \\ &= \alpha \int_0^t x_1(s)\, ds + \beta \int_0^t x_2(s)\, ds = \alpha T x_1 + \beta T x_2.\end{aligned}$$

Furthermore, for each $x \in C[0,1]$ we have:

$$||Tx|| = \max_{t\in[0,1]} \left| \int_0^t x(s)\, ds \right| \leq \max_{t\in[0,1]} \int_0^t |x(s)|\, ds$$
$$\leq \max_{t\in[0,1]} \int_0^t ||x||\, ds = \max_{t\in[0,1]} t \cdot ||x|| = ||x||.$$

Hence T is bounded with $||T|| \leq 1$.

Now assume $y = Tx$ for an arbitrary vector $x \in C[0,1]$. This means that $y(t) = \int_0^t x(s)\, ds$. This relation implies that $y(t)$ is differentiable with respect to t and that

$$y'(t) = \frac{d}{dt} \int_0^t x(s)\, ds = x(t).$$

(Here if $t = 0$ we interprete $y'(t)$ as *right derivative* $y'(0) := \lim_{h\to 0+} \frac{y(0+h)-y(0)}{h}$, and if $t = 1$ we interprete $y'(t)$ as *left derivative* $y'(1) := \lim_{h\to 0-} \frac{y(1+h)-y(1)}{h}$.) This shows that x is uniquely determined once $y = Tx$ is known, i.e. T is injective. Hence T^{-1} exists, and the above formula shows that

$$(T^{-1}y)(t) = y'(t), \qquad \text{for all } \ y \in \mathcal{R}(T).$$

(With conventions for $t = 0, 1$ as before.)

One may also prove that $\mathcal{R}(T)$ consists exactly of those functions $y \in C[0,1]$ such that $y(0) = 0$ and $y'(t)$ exists and is continuous for all $t \in [0,1]$ (with conventions as above for $t = 0, 1$). However, we do *not* need this precise description of $\mathcal{R}(T)$ for the purpose of the present problem.

Given any $n \in \mathbb{Z}^+$ we let $x_n(t) = t^n$. Then $x_n \in C[0,1]$ and $||x_n|| = \max_{t\in[0,1]} |t^n| = 1$. We let

$$y_n(t) = Tx_n(t) = \int_0^t s^n\, ds = (n+1)^{-1} t^{n+1}.$$

Then $y_n \in \mathcal{R}(T)$ and $T^{-1} y_n = x_n$, and $||y_n|| = \max_{t\in[0,1]} |(n+1)^{-1} t^{n+1}| = (n+1)^{-1}$. It now follows that T^{-1} is not bounded, by the same argument as in the first solution: *If* T^{-1} were bounded, then there would exist a constant $c \geq 0$ such that

$$(*) \qquad\qquad ||T^{-1}(x)|| \leq c||x||, \qquad \forall x \in \mathcal{R}(T).$$

But then choose n as an integer *larger* than c; we then have $y_n \in \mathcal{R}(T)$ and $||T^{-1}(y_n)|| = 1 > c(n+1)^{-1} = c||y_n||$, which *contradicts* (*). Hence T^{-1} is *not* bounded.

2. Throughout this exercise we have to compute a lot of integrals of the form $J_n = \int_0^\infty t^n e^{-t}$, where $n \geq 0$ is an integer. This can be done by repeated integration by parts: Note that $J_0 = 1$, and for $n \geq 1$ we have

$$J_n = \int_0^\infty t^n e^{-t} = \left[t^n(-e^{-t})\right]_0^\infty - \int_0^\infty nt^{n-1}(-e^{-t})\,dt$$
$$= 0 + n\int_0^\infty nt^{n-1}e^{-t}\,dt = n \cdot J_{n-1}.$$

Hence for $n \geq 1$:

$$J_n = n \cdot J_{n-1} = n(n-1) \cdot J_{n-2} = ... = n! \cdot J_0 = n!$$

This formula is also true for $n = 0$. From this we obtain the following general formula in $L^2[0, +\infty]$:

$$\langle t^m e^{-t/2}, t^n e^{-t/2}\rangle = \int_0^\infty t^m e^{-t/2} \cdot \overline{t^n e^{-t/2}}\,dt = \int_0^\infty t^{m+n}e^{-t}\,dt J_{m+n} = (m+n)!$$

We will use this repeatedly below.

Note that x_1, x_2, x_3 are linearly independent. We now apply the Gram-Schmidt orthonormalization process to x_1, x_2, x_3, see pp. 157-158 in Kreyszig's book. First:

$$||x_1||^2 = \langle t^2 e^{-t/2}, t^2 e^{-t/2}\rangle = 4! = 24,$$

and hence

$$e_1 = \frac{1}{||x_1||} \cdot x_1 = \frac{1}{\sqrt{24}} \cdot t^2 e^{-t/2}.$$

Next $\langle x_2, e_1\rangle = \frac{1}{\sqrt{24}} \cdot \langle t^2 e^{-t}, te^{-t/2}\rangle = \frac{3!}{\sqrt{24}} = \sqrt{\frac{3}{2}}$, and hence, using the same notation as in the book:

$$v_2 = x_2 - \langle x_2, e_1\rangle e_1 = te^{-t/2} - \frac{3!}{\sqrt{24}}\frac{1}{\sqrt{24}} \cdot t^2 e^{-t/2} = te^{-t/2} - \frac{1}{4} \cdot t^2 e^{-t/2};$$

$$||v_2||^2 = \langle v_2, v_2\rangle = \langle te^{-t/2} - \frac{1}{4} \cdot t^2 e^{-t/2}, te^{-t/2} - \frac{1}{4} \cdot t^2 e^{-t/2}\rangle$$
$$= \langle te^{-t/2}, te^{-t/2}\rangle - \frac{1}{2}\langle te^{-t/2}, t^2 e^{-t/2}\rangle + \frac{1}{16}\langle t^2 e^{-t/2}, t^2 e^{-t/2}\rangle = 2! - \frac{3!}{2} + \frac{4!}{16} = \frac{1}{2};$$

$$e_2 = \frac{1}{||v_2||}v_2 = \sqrt{2}v_2 = \sqrt{2}\left(te^{-t/2} - \frac{1}{4} \cdot t^2 e^{-t/2}\right).$$

Finally,

$$\langle x_3, e_1\rangle = \frac{1}{\sqrt{24}} \cdot \langle e^{-t/2}, t^2 e^{-t/2}\rangle = \frac{2}{\sqrt{24}} = \frac{1}{\sqrt{6}};$$

$$\langle x_3, e_2\rangle = \langle e^{-t/2}, \sqrt{2}te^{-t/2} - \frac{\sqrt{2}}{4} \cdot t^2 e^{-t/2}\rangle = \sqrt{2} \cdot 1! - \frac{\sqrt{2}}{4} \cdot 2! = \frac{1}{\sqrt{2}},$$

and hence:

$$
\begin{aligned}
v_3 &= x_3 - \langle x_3, e_1\rangle e_1 - \langle x_3, e_2\rangle e_2 \\
&= e^{-t/2} - \frac{1}{\sqrt{6}} \cdot \frac{1}{\sqrt{24}} \cdot t^2 e^{-t/2} - \frac{1}{\sqrt{2}} \cdot \left(\sqrt{2} t e^{-t/2} - \frac{\sqrt{2}}{4} \cdot t^2 e^{-t/2} \right) \\
&= \frac{1}{6} t^2 e^{-t/2} - t e^{-t/2} + e^{-t/2}; \\
||v_3||^2 = \langle v_3, v_3\rangle &= \int_0^\infty \left| \frac{1}{6} t^2 e^{-t/2} - t e^{-t/2} + e^{-t/2} \right|^2 \, dt \\
&= \int_0^\infty \left(\frac{1}{36} t^4 e^{-t} - \frac{1}{3} t^3 e^{-t} + \frac{4}{3} t^2 e^{-t} - 2 t e^{-t} + e^{-t} \right) \, dt \\
&= \frac{4!}{36} - \frac{3!}{3} + \frac{4 \cdot 2!}{3} - 2 + 1 = \frac{1}{3}; \\
e_3 = \frac{1}{||v_3||} v_3 &= \sqrt{3} \left(\frac{1}{6} t^2 e^{-t/2} - t e^{-t/2} + e^{-t/2} \right).
\end{aligned}
$$

Answer: The orthonormalized basis is

$$
\begin{aligned}
e_1 &= \frac{1}{\sqrt{24}} t^2 e^{-t/2} = \frac{\sqrt{6}}{12} t^2 e^{-t/2}, \\
e_2 &= \sqrt{2} \left(t e^{-t/2} - \frac{1}{4} \cdot t^2 e^{-t/2} \right) = \frac{\sqrt{2}}{4} (4t - t^2) e^{-t/2}, \\
e_3 &= \sqrt{3} \left(\frac{1}{6} t^2 e^{-t/2} - t e^{-t/2} + e^{-t/2} \right) = \frac{\sqrt{3}}{6} (6 - 6t + t^2) e^{-t/2}.
\end{aligned}
$$

Alternative solution. Just for fun, let us deduce the same result from the facts given in §3.7 (this section is *not* part of the course content, but I have recommended that you read it anyway). From §3.7-3 we learn that the following vectors are orthonormal in $L^2[0, \infty]$:

$$
f_1 = e^{-t/2}; \quad f_2 = (1 - t) e^{-t/2}; \quad f_3 = (1 - 2t + \tfrac{1}{2}) e^{-t/2}.
$$

(In fact, we learn from §3.7-3 that f_1, f_2, f_3 are obtained if our vectors x_1, x_2, x_3 are orthonormalized in the order x_3, x_2, x_1.) Now we see by inspection:

$$
x_1 = 2f_1 - 4f_2 + 2f_3; \quad x_2 = f_1 - f_2; \quad x_3 = f_1.
$$

Now it is very easy to apply the Gram-Schmidt orthonormalization process, using the fact that f_1, f_2, f_3 are orthonormal:

$$
\begin{aligned}
||x_1|| &= \sqrt{2^2+4^2+2^2} = \sqrt{24};\\
e_1 &= \frac{1}{||x_1||}x_1 = \tfrac{1}{\sqrt{24}}(2f_1 - 4f_2 + 2f_3) = \tfrac{1}{\sqrt{6}}(f_1 - 2f_2 + f_3);\\
v_2 &= x_2 - \langle x_2, e_1\rangle e_1\\
&= (f_1 - f_2) - \Big\langle f_1 - f_2, \tfrac{1}{\sqrt{6}}(f_1 - 2f_2 + f_3)\Big\rangle \cdot \tfrac{1}{\sqrt{6}}(f_1 - 2f_2 + f_3)\\
&= f_1 - f_2 - \tfrac{1}{6}(1+2)(2f_1 - 4f_2 + 2f_3) = \tfrac{1}{2}f_1 - \tfrac{1}{2}f_3;\\
e_2 &= \frac{1}{||v_2||}v_2 = \frac{1}{\sqrt{\frac{1}{2}^2 + \frac{1}{2}^2}}(\tfrac{1}{2}f_1 - \tfrac{1}{2}f_3) = \tfrac{\sqrt{2}}{2}(f_1 - f_3);\\
v_3 &= x_3 - \langle x_3, e_1\rangle e_1 - \langle x_3, e_2\rangle e_2\\
&= f_1 - \Big\langle f_1, \tfrac{1}{\sqrt{6}}(f_1 - 2f_2 + f_3)\Big\rangle \cdot \tfrac{1}{\sqrt{6}}(f_1 - 2f_2 + f_3)\\
&\qquad - \langle f_1, \tfrac{\sqrt{2}}{2}(f_1 - f_3)\rangle \cdot \tfrac{\sqrt{2}}{2}(f_1 - f_3)\\
&= f_1 - \frac{1}{6}(f_1 - 2f_2 + f_3) - \frac{1}{2}(f_1 - f_3) = \frac{1}{3}(f_1 + f_2 + f_3);\\
e_3 &= \frac{1}{||v_3||}v_3 = \frac{1}{\sqrt{\frac{1}{3}^2 + \frac{1}{3}^2 + \frac{1}{3}^2}}\frac{1}{3}(f_1 + f_2 + f_3) = \frac{\sqrt{3}}{3}(f_1 + f_2 + f_3).
\end{aligned}
$$

Substituting the formulae for f_1, f_2, f_3 we check that we have obtained the same orthonormal vectors e_1, e_2, e_3 as in the first solution.

3. (a) Take $y \in H_2$. Take $x \in M_1$. Then $\langle T^*(y), x\rangle = \langle y, Tx\rangle = \langle y, 0\rangle = 0$ (the second equality holds because $x \in M_1 = \mathcal{N}(T)$). Hence we have proved that $\langle T^*(y), x\rangle = 0$ for all $x \in M_1$; this means that $T^*(y) \in M_1^\perp$. This holds for all $y \in H_2$, hence $T^*(H_2) \subset M_1^\perp$.

(b) Take $y \in [T(H_1)]^\perp$. Then $y \in H_2$ and we wish to prove that $T^*(y) = 0$. Note that for all $x \in H_2$ we have

$$\langle T^*(y), x\rangle = \langle y, Tx\rangle = 0,$$

where the last equality holds because $y \in [T(H_1)]^\perp$ and $Tx \in T(H_1)$. Since $\langle T^*(y), x\rangle = 0$ holds for all $x \in H_2$ we have $T^*(y) = 0$ (by Lemma 3.8-2). Hence $y \in \mathcal{N}(T^*)$. This holds for all $y \in [T(H_1)]^\perp$, hence we have proved $[T(H_1)]^\perp \subset \mathcal{N}(T^*)$.

Remark: In fact we have $[T(H_1)]^\perp = \mathcal{N}(T^*)$ (cf. below).

(c) Take $x \in M_1$. Take $y \in H_2$. Then $\langle x, T^*(y)\rangle = \langle Tx, y\rangle = \langle 0, y\rangle = 0$. This is true for all $y \in H_2$; in other words $\langle x, z\rangle = 0$ for all

$z \in [T^*(H_2)]$. Hence $x \in [T^*(H_2)]^\perp$. This is true for all $x \in M_1$. Hence we have proved

$$(*) \qquad M_1 \subset [T^*(H_2)]^\perp.$$

Conversely, take $x \in [T^*(H_2)]^\perp$. Then $\langle x, T^*(y)\rangle = 0$ for all $y \in H_2$. Hence $\langle Tx, y\rangle = 0$ for all $y \in H_2$. Hence $Tx = 0$ (by Lemma 3.8-2). Hence $x \in \mathcal{N}(T) = M_1$. This is true for all $x \in [T^*(H_2)]^\perp$. Hence we have proved

$$(**) \qquad [T^*(H_2)]^\perp \subset M_1.$$

Together, (*) and (**) imply that $M_1 = [T^*(H_2)]^\perp$.

Alternative solution. We do the three parts in opposite order:
(c) We have

$$\begin{aligned} [T^*(H_2)]^\perp &=^1 \{x \in H_1 \mid \forall z \in T^*(H_2) : \langle x, z\rangle = 0\} \\ &=^2 \{x \in H_1 \mid \forall y \in H_2 : \langle x, T^*(y)\rangle = 0\} \\ &=^3 \{x \in H_1 \mid \forall y \in H_2 : \langle Tx, y\rangle = 0\} \\ &=^4 \{x \in H_1 \mid Tx = 0\} \\ &=^5 \mathcal{N}(T) = M_1. \end{aligned}$$

1. By definition of orthogonal complement.
2. By definition of $T^*(H_2)$.
3. By definition of T^*.
4. By Lemma 3.8-2 and the trivial fact that $\langle 0, y\rangle = 0$ for all $y \in H_2$.
5. By definition of $\mathcal{N}(T)$

(b) In (c) we proved that $[T^*(H_2)]^\perp = \mathcal{N}(T)$ holds for every bounded linear operator $T : H_1 \longrightarrow H_2$. If we apply this fact to the bounded linear operator $T^* : H_2 \rightarrow H_1$ we obtain $[T^{**}(H_1)]^\perp = \mathcal{N}(T^*)$. But $T^{**} = T$, hence $[T(H_1)]^\perp = \mathcal{N}(T^*)$. This is a *stronger* statement than what we had to prove in (b)!

(a) Since $[T^*(H_2)]^\perp = M_1$ (as proved in (c)), we have $[T^*(H_2)]^{\perp\perp} = M_1^\perp$. But we also know $A \subset A^{\perp\perp}$, for *any* subset $A \subset H_1$. In particular, $T^*(H_2) \subset [T^*(H_2)]^{\perp\perp} = M_1^\perp$.

4. (We assume $r > 0$, since the sphere $S(0; r)$ has only been defined for such r in the book.) Take $x_0 \in S(0; r)$. Then $||x_0|| = r > 0$, and thus $x_0 \neq 0$. Hence by Theorem 4.3-3 there exists some $f \in X'$ such that $||f|| = 1$ and $f(x_0) = ||x_0|| = r$. Let H be the hyperplane

$$H = \{x \in X \mid f(x) = r\}.$$

Then clearly $x_0 \in H$. Furthermore, for each $x \in \tilde{B}(0; r)$ we have $||x|| \leq r$ and hence $f(x) \leq |f(x)| \leq ||f|| \cdot ||x|| \leq 1 \cdot r = r$. (In the first inequality we used $f(x) \in \mathbb{R}$, since $K = \mathbb{R}$ in this problem.) Hence we have proved:

$$\tilde{B}(0;r) \subset \{x \in X \mid f(x) \leq r\}.$$

This means that $\tilde{B}(0;r)$ lies completely in the half space $\{x \in X \mid f(x) \leq r\}$, which is one of the two half spaces determined by H.

5. We first assume that $T : H \to H$ is a unitary operator. Let M be an arbitrary total orthonormal subset in H. Take $w_1, w_2 \in T(M)$. Then we have $w_1 = Tv_1$ and $w_2 = Tv_2$ for some $v_1, v_2 \in H$. Using the fact that T is unitary and that M is an orthonormal set, we get:

$$\begin{aligned} \langle w_1, w_2 \rangle &= \langle Tv_1, Tv_2 \rangle = \langle v_1, T^*Tv_2 \rangle = \langle v_1, T^{-1}Tv_2 \rangle \\ &= \langle v_1, v_2 \rangle = \begin{cases} 1 & \text{if } v_1 = v_2 \\ 0 & \text{if } v_1 \neq v_2. \end{cases} \end{aligned}$$

But T is a bijection since T is unitary; hence $v_1 = v_2 \iff Tv_1 = Tv_2 \iff w_1 = w_2$. Hence we have proved

$$\langle w_1, w_2 \rangle = \begin{cases} 1 & \text{if } w_1 = w_2 \\ 0 & \text{if } w_1 \neq w_2, \end{cases}$$

for all $w_1, w_2 \in T(M)$. Hence $T(M)$ *is an orthonormal set.*

Next we will use Theorem 3.6-2 to prove that $T(M)$ is total. Let $x \in H$ be an arbitrary vector such that $x \perp T(M)$. Then $\langle x, Tv \rangle = 0$ for all $v \in M$, and hence since T is unitary, $\langle T^{-1}x, v \rangle = 0$ for all $v \in M$. By Theorem 3.6-2(a) this implies that $T^{-1}x = 0$, since M is total in H. But $T^{-1}x = 0$ implies $x = 0$ (since T^{-1} is always injective if it exists). Hence we have proved that

$$\forall x \in H : \quad x \perp T(M) \Longrightarrow x = 0.$$

Hence by Theorem 3.6-2(b) (which is applicable since H is a Hilbert space), $T(M)$ is total in H.

Hence if T is unitary then for every total orthonormal set M in H we have proved that $T(M)$ is a total orthonormal set in H.

Conversely, assume that $T : H \to H$ be a linear operator such that $T(M)$ is a total orthonormal set in H for each total orthonormal set M. Let us first prove that T *is bounded.* Given a fixed vector $x \in H$ with $||x|| = 1$, let $Y = \text{Span}\{x\}$; this is a closed subspace of H by Theorem 2.4-3 and hence by Theorem 3.3-4 we have $H = Y \oplus (Y^\perp)$. Also $Y^\perp$ is closed subspace of H and hence $Y^\perp$ is a Hilbert space in itself. Hence by p.168 (middle) (cf. Theorem 4.1-8) there exists a total orthonormal subset $M_1 \subset Y^\perp$. Now let $M = \{x\} \cup M_1$; this is clearly an orthonormal set since $||x|| = 1$ and M_1 is orthogonal to $Y = \text{Span}\{x\}$ and hence to x.

We also have $\overline{\text{Span}(M)} \supset \overline{\text{Span}(M_1)} = Y^\perp$ and $\overline{\text{Span}(M)} \supset \overline{\text{Span}\{x\}} = Y$, and hence $\overline{\text{Span}(M)}$ contains every vector in $Y \oplus (Y^\perp) = H$. Hence M is a total orthonormal set in H. By our assumption, this implies that $T(M)$ is a total orthonormal set in H, and since $x \in M$ we get in particular $||T(x)|| = 1$.

We have thus proved that $||T(x)|| = 1$ for every $x \in H$ with $||x|| = 1$. Hence T *is bounded* and $||T|| = 1$. It now also follows directly that

$$(*) \qquad ||T(y)|| = ||y||, \qquad \forall y \in H.$$

(Proof: If $y = 0$ then trivially $||T(y)|| = ||0|| = 0$. Now assume $y \neq 0$. Then $y = ||y|| \cdot x$ where $x = ||y||^{-1} \cdot y \in H$ and $||x|| = 1$, hence by what we have showed, $||T(x)|| = 1$, and thus $||T(y)|| = ||T(||y|| \cdot x)|| = ||y|| \cdot ||T(x)|| = ||y||$.)

From (*) one deduces directly that T *is injective*. (This is something which I have pointed out in a lecture. The proof is as follows: Assume $T(y_1) = T(y_2)$. Then $T(y_1 - y_2) = 0$, thus $||T(y_2 - y_1)|| = 0$, and hence by (*), $||y_2 - y_1|| = 0$, i.e. $y_1 = y_2$. This shows that T is injective.)

We next prove that T *is surjective*. Let M be an arbitrary total orthonormal set in H. By our assumption $T(M)$ is a total orthonormal set, and hence $\overline{\text{Span}(T(M))} = H$. But $T(M) \subset T(H)$, and $T(H)$ is a subspace of H, and thus $\text{Span}(T(M)) \subset T(H)$ and $H = \overline{\text{Span}(T(M))} \subset \overline{T(H)}$. Hence $\overline{T(H)} = H$. Now fix an arbitrary element $y \in H$; we wish to construct a vector $x \in H$ such that $T(x) = y$. Since $y \in H = \overline{T(H)}$ there is a sequence $y_1, y_2, ...$ in $T(H)$ such that $y_j \to y$. Since $y_j \in T(H)$ we may write $y_j = T(x_j)$ for some $x_j \in H$. Using now (*) and then Theorem 1.4-5 we get

$$||x_j - x_k|| = ||T(x_j - x_k)|| = ||y_j - y_k|| \to 0 \qquad \text{as } j, k \to \infty.$$

Hence $x_1, x_2, ...$ is a Cauchy sequence in H, and since H is a Hilbert space (i.e. complete) there is a vector $x \in H$ such that $x_j \to x$. Since T is bounded (and hence continuous) we now have

$$T(x) = T(\lim_{j \to \infty} x_j) = \lim_{j \to \infty} T(x_j) = \lim_{j \to \infty} y_j = y.$$

Hence for each $y \in H$ there is some $x \in H$ such that $T(x) = y$. This proves that T *is surjective*.

Next, by using *polarization* (p.134 (9), (10)), one deduces from (*) that

$$(**) \qquad \langle Tx, Ty \rangle = \langle x, y \rangle, \qquad \forall x, y \in H.$$

(Explanation: Formulas p.134 (9), (10) show that the inner product in H may be expressed completely in terms of the norm; hence since

(*) shows that T preserves the norm, T must also preserve the inner product! If writes out the computation it it looks as follows. If $K = \mathbb{R}$:

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

If $K = \mathbb{C}$:

$$\text{Re } \langle Tx, Ty\rangle = \tfrac{1}{4}(||Tx + Ty||^2 - ||Tx - Ty||^2) = \tfrac{1}{4}(||T(x + y)||^2 - ||T(x - y)||^2)$$
$$= \tfrac{1}{4}(||x + y||^2 - ||x - y||^2) = \text{Re } \langle x, y\rangle$$

and

$$\text{Im } \langle Tx, Ty\rangle = \tfrac{1}{4}(||Tx + iTy||^2 - ||Tx - iTy||^2) = \tfrac{1}{4}(||T(x + iy)||^2 - ||T(x - iy)||^2)$$
$$= \tfrac{1}{4}(||x + iy||^2 - ||x - iy||^2) = \text{Im } \langle x, y\rangle;$$

hence the numbers $\langle Tx, Ty\rangle$ and $\langle x, y\rangle$ have the same real part and the same imaginary part; hence $\langle Tx, Ty\rangle = \langle x, y\rangle$.)

Now note that (**) implies $\langle T^*Tx, y\rangle = \langle Tx, Ty\rangle = \langle x, y\rangle$ for all $x, y \in H$, hence by Lemma 3.8-2, $T^*Tx = x$ for all $x \in H$. Since T is bijective, this relation implies $T^{-1} = T^*$.

6. In fact we do not have to assume that Y is closed; hence from now on let Y be an *arbitrary* subspace of the normed space X.

We first check carefully that the various concepts introduced in the problem are well-defined: The annihilator Y^a is defined in problem 13, Section 2.10, and from that problem we know that Y^a *is a closed subspace of* X'. Hence X'/Y^a is a normed space by problem 14, Section 2.3. Finally we check that the map $A : X'/Y^a \to Y'$ is well-defined: Take any $f, g \in X'$ such that $f + Y^a = g + Y^a$. We then have to prove that $A(f + Y^a)$ and $A(g + Y^a)$ are defined to be the same thing, i.e. that $f_{|Y} = g_{|Y}$. But $f + Y^a = g + Y^a$ implies $f = g + h$ for some $h \in Y^a$, and hence for each $y \in Y$ we have $f(y) = g(y) + h(y) = g(y) + 0$. Hence $f_{|Y} = g_{|Y}$, as desired.

We now start our proof that A is an isomorphism of normed spaces. First of all, for any $f, g \in X'$ and any $\alpha, \beta \in K$ we have

$$A(\alpha(f + Y^a) + \beta(g + Y^a)) = A((\alpha f + \beta g) + Y^a) = (\alpha f + \beta g)_{|Y}$$
$$= \alpha f_{|Y} + \beta g_{|Y} = \alpha A(f + Y^a) + \beta A(g + Y^a).$$

(In the first equality we used the definition of addition and multiplication in X'/Y^a, see problem 14 in Section 2.1.) Hence A *is a linear operator.*

Next, let $f \in X'$ be given; we wish to prove $||f + Y^a|| = ||A(f + Y^a)||$, i.e. $||f + Y^a|| = ||f_{|Y}||$, By the definition of the norm on X'/Y^a (see problem 14, Section 2.3) we have

$$(*) \qquad ||f + Y^a|| = \inf_{g \in f + Y^a} ||g||.$$

Take any $g \in f + Y^a$. Then $g = f + h$ for some $h \in Y^a$, and hence for all $y \in Y$ we have $g(y) = f(y) + h(y) = f(y)$. Hence, using the fact $Y \subset X$:

$$||g|| = \sup_{x \in X - \{0\}} \frac{|g(x)|}{||x||} \geq \sup_{y \in Y - \{0\}} \frac{|g(y)|}{||y||} = \sup_{y \in Y - \{0\}} \frac{|f(y)|}{||y||} = ||f_{|Y}||.$$

Since this is true for all $g \in f + Y^a$ we have by (*):

$$(**) \qquad ||f + Y^a|| \geq ||f_{|Y}||.$$

On the other hand, by Hahn-Banach's Theorem 4.3-2 (applied to the subspace $Y \subset X$ and the bounded linear functional $f_{|Y}$ on Y), there exists some $g \in X'$ such that $g_{|Y} = f_{|Y}$ (i.e. g is an extension of $f_{|Y}$) and $||g|| = ||f_{|Y}||$. Let $h = g - f \in X'$. Then for all $y \in Y$ we have $h(y) = g(y) - f(y) = 0$, since $g_{|Y} = f_{|Y}$. Thus $h \in Y^a$. Hence we have $g = f + h$ and $h \in Y^a$; hence $g \in f + Y^a$. Hence by (*):

$$(***) \qquad ||f + Y^a|| \leq ||g|| = ||f_{|Y}||.$$

By (**) and (***) we have finally proved

$$||f + Y^a|| = ||A(f + Y^a)|| = ||f_{|Y}||,$$

i.e. the linear operator $A : X'/Y^a \to Y'$ is norm preserving.

Since A is norm preserving A *is injective* (as we also mentioned in problem 5). Finally, we prove that A is surjective: Let g be an arbitrary element in Y'. Then by Hahn-Banach's Theorem 4.3-2 there exists some $f \in X'$ such that $f_{|Y} = g$ and $||f|| = ||g||$. Now $f + Y^a \in X'/Y^a$ and $A(f + Y^a) = f_{|Y} = g$. This proves that A *is surjective.*

We have now proved that A is a bijective and norm preserving linear map. In other words, A *is an isomorphism of normed spaces.*