Orthogonal projections and the spectral theorem

Def 1) We say that $T \in \mathcal{L}(V)$ is the projection on W_1 along W_2, where W_1, W_2 are subspaces of V s.t. $V = W_1 \oplus W_2$, if whenever $x = x_1 + x_2$, with $x_i \in W_i$, then $T(x) = x_1$.
2) By a H/W exercise, we know that then $R(T) = W_1$ and $N(T) = W_2$.
3) We say that $T \in \mathcal{L}(V)$ is a projection if there exist W_1 and W_2 satisfying the definition above for T.
4) By H/W, T is a projection if and only if $T = T^2$.

Because $V = W_1 \oplus W_2 = W_1 \oplus W_3$ does not imply that $W_2 = W_3$, we see that W_1 alone does not determine T uniquely. But it does for orthogonal projections, as defined below.

Def Let V be an I.P.S., and let $T \in \mathcal{L}(V)$ be a projection. We say that T is an orthogonal projection if $R(T)^\perp = N(T)$ and $N(T)^\perp = R(T)$.

Rem From 115A, we know that if V is finite dimensional and W is a subspace of V, then $(W^\perp)^\perp = W$. Hence each of the two conditions in this definition already implies the other one: if $R(T)^\perp = N(T)$, then $R(T) = R(T)^{\perp\perp} = N(T)^\perp$, and vice versa.

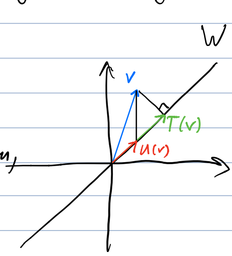
Rem This is not true if $\dim(V) = \infty$.
Consider the space $V = \ell^2$, i.e. $V = \{(a_i)_{i \in \mathbb{N}} : (a_i) \text{ is an infinite sequence of real numbers and } \sum_{i \in \mathbb{N}} a_i^2 < \infty\}$.
It is a R.I.P.S. with the following operations:
$(a_i)_{i\in\mathbb{N}} + (b_i)_{i\in\mathbb{N}} = (a_i + b_i)_{i\in\mathbb{N}} \quad \forall (a_i), (b_j) \in \ell^2, \; c \in \mathbb{R}$
$c \cdot (a_i)_{i\in\mathbb{N}} = (c \cdot a_i)_{i\in\mathbb{N}}$
$\langle (a_i)_{i\in\mathbb{N}}, (b_i)_{i\in\mathbb{N}} \rangle = \sum_{i=1}^{\infty} a_i b_i$.

Note that $\dim(V) = \infty$. Consider the subspace
$W = \{(a_i) \in \ell^2 : \exists j \text{ s.t. } a_i = 0 \text{ for all } i > j\}$.
Then we have:
- $W^\perp = \{0\}$ (Exercise! Hint: use that W is dense in V, i.e. $\forall \varepsilon \in \mathbb{R}_{>0} \; \forall x \in V \; \exists y \in W \; \|x - y\| < \varepsilon$.
- $W^{\perp\perp} = \{0\}^\perp = V$.

But $W \neq V$ (e.g. the sequence $(\frac{1}{2^i})_{i\in\mathbb{N}} \in V$ as $\sum_{i=1}^{\infty} \frac{1}{2^i} < \infty$, but $(\frac{1}{2^i})_{i\in\mathbb{N}} \notin W$ as $\frac{1}{2^i} \neq 0 \; \forall i$).
So $W \neq W^{\perp\perp}$.

Let W be a subspace of V, $\dim(W) < \infty$.
By Thm 6.6, every $y \in V$ can be written as:
$y = u + z$ for unique $u \in W$ and $z \in W^\perp$.
Define the map $T : V \to V$ by $T(y) = u$.
We see from definition above that T is an orthogonal projection on W.
Moreover, there exists exactly one orthogonal projection on W:
if T, U are orthogonal projections on W, then $R(T) = W = R(U)$; hence $N(T) = R(T)^\perp = R(U)^\perp = N(U)$; then for any $y \in V$ we have $y = u + z$ for unique $u \in R(T) = R(U)$ and $z \in N(T) = N(U)$. By def. of projection, $T(y) = u = U(y)$.

Def We call T the orthogonal projection of V on W.



Ex Difference between projection on W and orthogonal projection on W.
Let $V = \mathbb{R}^2$, $W = \mathrm{Span}\{(1,1)\}$. Let U, T be as in the picture. Then T is an orthogonal projection, but U is a different projection on W.
Note: $v - T(v) \in W^\perp$, $v - U(v) \notin W^\perp$.

We have an algebraic description of orthogonal projections:

Thm 6.24 Let V be an I.P.S. and $T \in \mathcal{L}(V)$. Then:
T is an orthogonal projection (on some subspace W) $\iff$ T has an adjoint T^* and $T^2 = T = T^*$.

Proof

(we are not assuming $\dim < \infty$).

$\Rightarrow$ Supp. T is an orthogonal projection.
As $T^2 = T$ because T is a projection, only need to show that T^* exists and $T = T^*$.
By def., $V = R(T) \oplus N(T)$ and $R(T)^\perp = N(T)$.
Let $x, y \in V$, then $x = x_1 + x_2$ and $y = y_1 + y_2$ for some $x_1, y_1 \in R(T)$, $x_2, y_2 \in N(T)$. Then:
$\langle x, T(y)\rangle = \langle x_1 + x_2, y_1\rangle = \langle x_1, y_1\rangle + \underbrace{\langle x_2, y_1\rangle}_{=0 \text{ as } y_1 \in R(T),\ x_2 \in N(T) = R(T)^\perp} = \langle x_1, y_1\rangle$

And $\langle T(x), y\rangle = \langle x_1, y_1 + y_2\rangle = \langle x_1, y_1\rangle + \underbrace{\langle x_1, y_2\rangle}_{=0} = \langle x_1, y_1\rangle$.

So $\langle x, T(y)\rangle = \langle T(x), y\rangle$ for all $x, y \in V$, thus T^* exists and $T = T^*$.

$\Leftarrow$ Suppose $T^2 = T = T^*$. Then T is a projection by H/W, and we must show $R(T) = N(T)^\perp$ and $R(T)^\perp = N(T)$.
Let $x \in R(T)$ and $y \in N(T)$. Then $x = T(x) = T^*(x)$, so
$\langle x, y\rangle = \langle T^*(x), y\rangle = \langle x, T(y)\rangle = \langle x, 0\rangle = 0 \Rightarrow x \in N(T)^\perp$. So $R(T) \subseteq N(T)^\perp$.
Let $y \in N(T)^\perp$. Then
$\|y - T(y)\|^2 \overset{\text{def}}{=} \langle y - T(y), y - T(y)\rangle = \langle y, y - T(y)\rangle - \langle T(y), y - T(y)\rangle$.
We have $T(y - T(y)) = T(y) - T^2(y) = T(y) - T(y) = 0$, so $y - T(y) \in N(T)$, hence $\langle y, y - T(y)\rangle = 0$.
Also $T^*(y - T(y)) = T^*(y) - T^*T(y) \overset{\text{by ass.}}{=} T^*(y) - T^*(y) = 0$, hence
$\langle T(y), y - T(y)\rangle = \langle y, T^*(y - T(y))\rangle = \langle y, 0\rangle = 0$.
Combining, $\|y - T(y)\|^2 = 0$, hence $y = T(y) \in R(T)$. So $N(T)^\perp \subseteq R(T)$, so $R(T) = N(T)^\perp$.

Then $R(T)^\perp = N(T)^{\perp\perp} \supseteq N(T)$ by H/W.
Supp. $x \in R(T)^\perp$. For any $y \in V$, we have $\langle T(x), y\rangle = \langle x, T^*(y)\rangle = \langle x, T(y)\rangle = 0$.
So $T(x) = 0$, thus $x \in N(T)$.
Hence $R(T)^\perp = N(T)$.

Rem 1) Let V be an I.P.S., $\dim(V) < \infty$, and W is a subspace of V. Thm 6.7
Let T be the orthogonal projection of V on W.
Let $\{v_1, \dots, v_k\}$ be an orthonormal basis for W, and extend it to an orthonormal basis $\{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$ for W. Then $[T]_\beta$ is of the form

$$\begin{pmatrix} I_k & O_1 \\ O_2 & O_3 \end{pmatrix}$$
with O_1, O_2, O_3 zero matrices.

2) If U is any projection on W, we may choose a basis γ for V s.t. $[U]_\gamma$ has the form above, but γ is <u>not necessarily orthonormal</u>.

Thm 6.25 (The Spectral Theorem)
Supp. $T \in \mathcal{L}(V)$, V an I.P.S. over F, $\dim(V) < \infty$.
Let $\lambda_1, \dots, \lambda_k \in F$ be the distinct e.val's of T.
Assume T is normal if $F = \mathbb{C}$ or self-adjoint if $F = \mathbb{R}$.
For each i, $1 \le i \le k$, let $W_i = E_{\lambda_i}$ be the e.space of T corresponding to λ_i.
Let T_i be the orthogonal projection of V on W_i.
Then the following holds:

a) $V = W_1 \oplus \dots \oplus W_k$.

b) Let $W_i' = \bigoplus_{j \neq i} W_j$. Then $W_i^\perp = W_i'$.

c) $T_i T_j = \delta_{ij} T_i$ for $1 \le i, j \le k$.

d) $I = T_1 + \dots + T_k$. — the resolution of the identity operator.

e) $T = \lambda_1 T_1 + \dots + \lambda_k T_k$. — the spectral decomposition of T.

The set $\{\lambda_1, \dots, \lambda_k\}$ is called the spectrum of T.

Proof

a) By Thm 6.16 if $F = \mathbb{C}$ and Thm 6.17 if $F = \mathbb{R}$, T is diagz.
Hence $V = W_1 \oplus \dots \oplus W_k$ by Thm 5.11.

b) If $x \in W_i$ and $y \in W_j$ for some $i \neq j$, then $\langle x, y \rangle = 0$ (by Thm 6.15(d)).
Hence $W_i' \subseteq W_i^\perp$.
By (a) we have $\dim(W_i') = \sum_{j \neq i} \dim(W_j) = \dim(V) - \dim(W_i)$.

On the other hand, $\dim(W_i^\perp) = \dim(V) - \dim(W_i)$ by Thm 6.7(c).
Hence $\dim(W_i') = \dim(W_i^\perp)$, so $W_i' = W_i^\perp$.

c) H/W 6.

d) As T_i is the orthogonal projection of V on W_i, from (b) we have:
$N(T_i) = R(T_i)^\perp = W_i^\perp = W_i'$.
By a), every $x \in V$ can be written as
$x = x_1 + \dots + x_k$ for some $x_i \in W_i$.
Then $T_i(x) = T_i(x_1) + \dots + T_i(x_i) + \dots + T_i(x_k) = T_i(x_i)$ since $\sum_{i \neq j} x_j \in W_i' = N(T_i)$, so $T(\sum_{i \neq j} x_j) = 0$.

def. of proj. $= x_i \in W_i$

So $x = T_1(x) + \dots + T_k(x)$ for all $x \in V$, so $I = T_1 + \dots + T_k$.

e) For any $x \in V$ we have $x = x_1 + \dots + x_k$ for some $x_i \in W_i$. Then
$T(x) = T(x_1) + \dots + T(x_k) = \lambda_1 x_1 + \dots + \lambda_k x_k$ (as $W_i = E_{\lambda_i}$)
$= \lambda_1 T_1(x) + \dots + \lambda_k T_k(x)$ (by (d))
$= (\lambda_1 T_1 + \dots + \lambda_k T_k)(x)$.

Rem Let β be the union of orthonormal bases of W_i's, let $m_i = \dim(W_i)$.
Then $[T]_\beta$ has the form

$$\begin{pmatrix} \boxed{\lambda_1 I_{m_1}} & & & 0 \\ & \boxed{\lambda_2 I_{m_2}} & & \\ & & \ddots & \\ 0 & & & \boxed{\lambda_k I_{m_k}} \end{pmatrix}$$

– a diagonal matrix whose entries on the diagonal are the e.val's λ_i, each repeated m_i times.

Lemma1 If $\lambda_1 T_1 + \dots + \lambda_k T_k$ is the spectral decomposition of T, then
$g(T) = g(\lambda_1) T_1 + \dots + g(\lambda_k) T_k$ for any polynomial g.

Proof H/W 6.

We obtain several corollaries.
Assume that $T \in L(V)$, V is an I.P.S. over a field F, $\dim(V) < \infty$.

Cor 1 If $F = \mathbb{C}$, then T is normal $\Leftrightarrow T^* = g(T)$ for some poly. $g(t)$ over F.

Proof $\Rightarrow$ Supp. T is normal.
Let $T = \lambda_1 T_1 + \dots + \lambda_k T_k$ – spectral decomp. of T, by Thm 6.25.
Then $T^* = (\lambda_1 T_1 + \dots + \lambda_k T_k)^* = \bar{\lambda}_1 T_1 + \dots + \bar{\lambda}_k T_k$ – using Thm 6.11, and $T_i = T_i^*$ by Thm 6.24.

Fact (Lagrange Interpolation) For any field F and any scalars $c_0, \dots, c_n \in F$ with $c_i \neq c_j$ for $i \neq j$, and any scalars $b_0, \dots, b_n \in F$ there exists a <u>unique</u> poly $g(t) \in P_n(F)$ s.t. $g(c_i) = b_i$ for all $i = 0, \dots, n$.

Using this fact, let g be s.t. $g(\lambda_i) = \bar{\lambda}_i \ \forall i$ (recall that λ_i are <u>distinct</u> e.val's). Then

$$g(T) \underset{\uparrow}{=} g(\lambda_1)T_1 + \dots + g(\lambda_k)T_k = \bar{\lambda}_1 T_1 + \dots + \bar{\lambda}_k T_k = T^* \quad \text{by the calculation above.}$$

by Lemma 1

$\Leftarrow$ Assume $T^* = g(T)$ for some poly $\overset{a_0 + a_1 t + \dots + a_n t^n}{g}$. Then, using linearity of T,

$$TT^* = T(a_0 + a_1 T + \dots + a_n T^n) = a_0 T + a_1 T^2 + \dots + a_n T^{n+1} = (a_0 + a_1 T + \dots + a_n T^n) T = T^* T,$$

so T is normal.

Cor 2 If $F = \mathbb{C}$, then T is unitary $\Leftrightarrow$ T is normal and $|\lambda| = 1$ for every e.val. λ of T.

Proof

$\Rightarrow$ If T is unitary, then T is normal and every e.val of T has abs. value 1 by Cor 2 of Thm 6.18.

$\Leftarrow$ Let $T = \lambda_1 T_1 + \dots + \lambda_k T_k$ be the spectral decomp. of T.

If $|\lambda| = 1$ for every e.val. λ of T, then by Spectral Theorem (c):

$TT^* = (\lambda_1 T_1 + \dots + \lambda_k T_k)(\bar{\lambda}_1 T_1 + \dots + \bar{\lambda}_k T_k)$ — as in Cor 1

$= |\lambda_1|^2 T_1 + \dots + |\lambda_k|^2 T_k$ — as $T_i T_j = \delta_{ij} T_i$ for $1 \leq i,j \leq k$ by Thm 6.25

$= T_1 + \dots + T_k$ — as $|\lambda_i| = 1$ for all i by assumpt.

$= I$ — by Thm 6.25 (d).

Hence T is unitary.

Cor 3 If $F = \mathbb{C}$ and T is normal, then T is self-adjoint $\Leftrightarrow$ every e.val. of T is real.

Proof

$\Leftarrow$ Let $T = \lambda_1 T_1 + \dots + \lambda_k T_k$ be the spectral decomp. of T.

Supp. every e.val. of T is real. Then

$T^* = \bar{\lambda}_1 T_1 + \dots + \bar{\lambda}_k T_k = \lambda_1 T_1 + \dots + \lambda_k T_k = T$.

$\Rightarrow$ Proved in the lemma before Thm 6.17.

Cor 4 Let T be as in the spectral thm, with spectral decomp. $T = \lambda_1 T_1 + \dots + \lambda_k T_k$. Then each T_j is a polynomial in T.

Proof

By Lagrange Interpolation, for each $1 \leq j \leq k$, there exists a poly. g_j s.t. $g_j(\lambda_i) = \delta_{ij}$ for all $1 \leq i \leq k$.

Then:

$$g_j(T) = g_j(\lambda_1) T_1 + \dots + g_j(\lambda_k) T_k = \delta_{1j} T_1 + \dots + \delta_{kj} T_k = T_j.$$