10.12 Some proofs

10.12.1 Lengths and reconstruction

Theorem 10.13. *Let V be a vector space over a field $\mathbb{F}$ and let $\langle,\rangle\colon V\times V\to\mathbb{F}$ be a bilinear form. Let $\|\ \|^2\colon V\to\mathbb{F}$ be the quadratic form associated to $\langle,\rangle$.*

(a) (Parallelogram property) If $x,y\in V$ then

$$\|x+y\|^2+\|x-y\|^2=2\|x\|^2+2\|y\|^2.$$

(b) (Pythagorean theorem) If $x,y\in V$ and $\langle x,y\rangle=0$ and $\langle y,x\rangle=0$ then

$$\|x\|^2+\|y\|^2=\|x+y\|^2.$$

(c) (Reconstruction) Assume that $\langle,\rangle$ is symmetric and that $2\neq 0$ in $\mathbb{F}$. Let $x,y\in V$. Then

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

Proof.
(a) Assume $x,y\in V$. Then

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

(b) Assume $x,y\in V$ and $\langle x,y\rangle=0$ and $\langle y,x\rangle=0$. Then

$$\begin{aligned}\|x+y\|^2&=\langle x+y,+x+y\rangle=\langle x,x\rangle+\langle x,y\rangle+\langle y,x\rangle+\langle y,y\rangle\\&=\|x\|^2+0+0+\|y\|^2=\|x\|^2+0+0+\|y\|^2.\end{aligned}$$

(c) Assume $x,y\in V$. Then

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

□

10.12.2 Cauchy-Schwarz

Theorem 10.14. *Let $\mathbb{F}$ be a field with an involution $\overline{}\colon\mathbb{F}\to\mathbb{F}$ such that the fixed field*

$$\mathbb{K}=\{a\in\mathbb{F}\mid a=\bar{a}\}\qquad\text{is an ordered field.}$$

For $a\in\mathbb{K}$ define

$$|a|^2=a\bar{a}.$$

Let V be an $\mathbb{K}$-vector space with a sesquilinear form $\langle,\rangle\colon V\times V\to\mathbb{F}$ such that

(a) If $x,y\in V$ then $\langle y,x\rangle=\overline{\langle x,y\rangle}$.

(b) If $x\in V$ then $\langle x,x\rangle\in\mathbb{K}_{\geq 0}$.

Let $\| \ \|\colon V \to \mathbb{F}$ be the corresponding quadratic form and assume that if $a \in \mathbb{K}_{\geq 0}$ then there exists a unique $c \in \mathbb{K}_{\geq 0}$ such that $c^2 = a$. Then

(c) (Cauchy-Schwarz) If $x, y \in V$ then $|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$.

(d) (Triangle inequality) If $x, y \in V$ then $\|x + y\| \leq \|x\| + \|y\|$.

Proof. (c) Let $x, y \in V$. If $x = 0$ then both sides of the Cauchy-Schwarz inequality are 0. Assume $x \neq 0$. The Gram-Schmidt process on the vectors (x, y) suggests the consideration of

$$u_1 = \frac{x}{\|x\|} \qquad \text{and} \quad u_2 = y - \frac{\langle y, x\rangle}{\langle x, x\rangle}x.$$

To avoid denominators, let $u = \langle x, x\rangle y - \langle y, x\rangle x$. Then

$$\begin{aligned} 0 \leq \langle u, u\rangle &= \big\langle \langle x, x\rangle y - \langle y, x\rangle x, \langle x, x\rangle y - \langle y, x\rangle x\big\rangle \\ &= \overline{\langle x, x\rangle}\langle x, x\rangle|\langle y, y\rangle - \langle x, x\rangle\overline{\langle y, x\rangle}\langle y, x\rangle - \langle y, x\rangle\overline{\langle x, x\rangle}\langle x, y\rangle + |\langle y, x\rangle|^2\langle x, x\rangle \\ &= \overline{\langle x, x\rangle}(\langle x, x\rangle\langle y, y\rangle - |\langle y, x\rangle|^2) \end{aligned}$$

Since $x \neq 0$ then $\langle x, x\rangle \in \mathbb{K}_{>0}$ and so $\overline{\langle x, x\rangle} = \langle x, x\rangle \in \mathbb{K}_{>0}$. Thus,

$$0 \leq \langle x, x\rangle\langle y, y\rangle - |\langle y, x\rangle|^2 \qquad \text{and so} \qquad |\langle y, x\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle.$$

Thus, BY SQUARES PROPOSITION, $|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$.
(d) Let $a \in \mathbb{F}$. Using that if $z \in \mathbb{F}$ then $|z|^2 = z\bar{z} \in \mathbb{K}_{\geq 0}$, then

$$|a + \bar{a}|^2 \leq |a + \bar{a}|^2 + |a - \bar{a}|^2 = (a + \bar{a})^2 - (a - \bar{a})^2 = 4a\bar{a} = 4|a|^2.$$

So $|a + \bar{a}| \leq 2|a|$. Also

$$\text{if } a + \bar{a} \in \mathbb{K}_{\leq 0} \text{ then } a + \bar{a} \leq 0 \leq |a + \bar{a}| \qquad \text{and} \qquad \text{if } a + \bar{a} \in \mathbb{K}_{\geq 0} \text{ then } a + \bar{a} = |a + \bar{a}|.$$

Combining these with $|a + \bar{a}| \leq 2|a|$ gives

$$a + \bar{a} \leq 2|a|.$$

Assume $x, y \in V$. Then

$$\begin{aligned} \|x + y\|^2 &= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &= \|x\|^2 + |y\|^2 + \langle x, y\rangle + \overline{\langle x, y\rangle} \\ &\leq \|x\|^2 + |y\|^2 + 2|\langle x, y\rangle| \\ &\leq \|x\|^2 + |y\|^2 + 2\|x| \cdot \|y\| \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$

Thus $\|x + y\| \leq \|x\| + \|y\|$. □

10.12.3 Dual basies

Proposition 10.15. *Let V be a vector space with a sesquilinear form $\langle,\rangle\colon V\times V\to\mathbb{F}$. Let $W\subseteq V$ be a subspace of V. Assume W is finite dimensional, that $(w_1,\ldots,w_k)$ is a basis of W and that G is the Gram matrix of $\langle,\rangle$ with respect to the basis $\{w_1,\ldots,w_k\}$. The following are equivalent:*

(a) A dual basis to $(w_1,\ldots,w_k)$ exists.

(b) G is invertible.

(c) $W\cap W^\perp=0$.

(d) The linear transformation

$$\begin{array}{rccc}\Psi_W\colon & W & \to & W^* \\ & v & \longmapsto & \varphi_v\end{array}\qquad \text{given by}\qquad \varphi_v(w)=\langle v,w\rangle,$$

is an isomorphism.

Proof.
(a) $\Rightarrow$ (b): Assume that $\{w^1,\ldots,w^k\}$ exists.
To show: G is invertible.
Define $H(\ell,i)\in\mathbb{F}$ by

$$w^i=\sum_{\ell=1}^k H(i,\ell)w_\ell.$$

Then

$$\delta_{ij}=\langle w^i,w_j\rangle=\sum_{\ell=1}^k H(i,\ell)\langle w_\ell,w_j\rangle=\sum_{\ell=1}^k H(i,\ell)G(\ell,j)=(HG)(i,j).$$

So $HG=1$, H is the inverse of G, and G is invertible.
(b) $\Rightarrow$ (a): Assume that G is invertible.
For $i\in\{1,\ldots,k\}$ define

$$w^i=\sum_{\ell=1}^k G^{-1}(i,\ell)w_\ell,\qquad \text{for } i\in\{1,\ldots,k\}.$$

Then

$$\langle w^i,w_j\rangle=\sum_{\ell=1}^k G^{-1}(i,\ell)\langle w_\ell,w_j\rangle=\sum_{\ell=1}^k G^{-1}(i,\ell)G(\ell,j)=(G^{-1}G)(i,j)=\delta_{ij}.$$

So $\{w^1,\ldots,w^k\}$ is a dual basis to $\{w_1,\ldots,w_k\}$.
(b) $\Rightarrow$ (c): Assume that G is invertible.
To show: $W\cap W^\perp=0$.
Let $w\in W\cap W^\perp$.
To show: $w=0$.
Write $w=c_1w_1+\cdots+c_kw_k$.
To show: If $j\in\{1,\ldots,k\}$ then $c_j=0$.
Since $w\in W^\perp$ then $\langle w,w_r\rangle=0$ for $r\in\{1,\ldots,k\}$ and

$$\begin{aligned}c_j&=\sum_{\ell=1}^n c_\ell\delta_{\ell j}=\sum_{\ell=1}^n c_\ell G(\ell,r)G^{-1}(r,j)\\&=\sum_{\ell=1}^k c_\ell\langle w_\ell,w_r\rangle G^{-1}(r,j)=\sum_{r=1}^k\langle w,w_r\rangle G^{-1}(r,j)=0.=\sum_{r=1}^k 0\cdot G^{-1}(r,j)=0.\end{aligned}$$

So $w = 0$.
(c) $\Rightarrow$ (b): Assume that $W \cap W^\perp = 0$.
To show: G is invertible.
To show: The rows of G are linearly independent.
To show: If $c_1, \ldots, c_k \in \mathbb{F}$ and $(c_1, \ldots, c_k)G = 0$ then $c_1 = 0,\ c_2 = 0, \ldots,\ c_k = 0$.
Assume $c_1, \ldots, c_k \in \mathbb{F}$ and $(c_1, \ldots, c_k)G = 0$.
To show: $c_1 = 0,\ c_2 = 0, \ldots,\ c_k = 0$.
Let $w = c_1 w_1 + \cdots + c_k w_k$.
If $i \in \{1, \ldots, k\}$ then, since $(c_1, \ldots, c_k)G = 0$,

$$0 = \sum_{\ell=1}^{k} c_\ell G(\ell, i) = \sum_{\ell=1}^{k} c_k \langle w_\ell, w_i\rangle = \langle c_1 w_1 + \cdots c_k w_k, w_i\rangle = \langle w, w_i\rangle.$$

So $w \in W^\perp$.
So $w \in W \cap W^\perp$.
So $w = 0$.
So $c_1 = 0,\ c_2 = 0,\ \ldots, c_k = 0$.
Thus the rows of G are linearly independent and G is invertible.
(c) $\Rightarrow$ (d): Assume that $W \cap W^\perp = 0$
To show: $\Psi_W \colon W \to W^*$ is an isomorphism.
To show: (ca) Ψ_W is injective.
(cb) Ψ_W is surjective.

(ca) Since $\ker(\Psi_W) = W \cap W^\perp$ then $\ker(\Psi_W) = 0$.
So Ψ_W is injective.
(cb) If $\{w_1, \ldots, w_k\}$ is a basis of W then defining $\varphi^i \colon W \to \mathbb{F}$ by

$$\text{if } c_1, \ldots, c_k \in \mathbb{F} \text{ then} \qquad \varphi^i(c_1 w_1 + \cdots + c_k w_k) = c_i,$$

produces a basis $\{\varphi^1, \ldots, \varphi^k\}$ of the dual space W^*.
So $\dim(W) = \dim(W^*)$.
Since Ψ_W is injective W is finite dimensional then $\dim(\mathrm{im}(\Psi_W)) = \dim(W) = \dim(W^*)$.
So $\mathrm{im}(\Psi_W) = W^*$ and ψ_W is surjective.

So Ψ_W is an isomorphism.
(d) $\Rightarrow$ (c): Assume that Ψ_W is an isomorphism.
So Ψ_W is injective.
So $\ker(\Psi_W) = 0$.
Since $\ker(\Psi_W) = W \cap W^\perp$ then $W \cap W^\perp = 0$. □

10.12.4 Nonisotropy

Proposition 10.16. *A sesquilinear form $\langle,\rangle \colon V \times V \to \mathbb{F}$ satisfies*

(no isotropic vectors condition) If $v \in V$ and $\langle v, v\rangle = 0$ then $v = 0$.

if and only if it satisfies

(no isotropic subspaces condition) If W is a subspace of V then $W \cap W^\perp = 0$.

Proof. $\Rightarrow$: Assume that if $v \in V$ and $\langle v, v\rangle = 0$ then $v = 0$.
To show: If W is a subspace of V then $W \cap W^\perp = 0$.
Assume W is a subspace of V.
To show: If $w \in W \cap W^\perp$ then $w = 0$.
Assume $w \in W \cap W^\perp$.
Then $\langle w, w\rangle = 0$.
So $w = 0$.

$\Leftarrow$: Assume that if W is a subspace of V then $W \cap W^\perp = 0$.
To show: If $v \in V$ and $\langle v, v\rangle = 0$ then $v = 0$.
Assume $v \in V$.
To show: If $v \neq 0$ then $\langle v, v\rangle \neq 0$.
Assume $v \neq 0$.
Let $W = \mathbb{F}v$, a one-dimensional subspace of V.
Since $\mathbb{F}v \cap (\mathbb{F}v)^\perp = 0$ then $v \notin (\mathbb{F}v)^\perp$.
So $\langle v, v\rangle \neq 0$. □

10.12.5 Orthogonal projections

Proposition 10.17. *(Characterization of orthogonal projection) Let $\mathbb{F}$ be a field and let V be an $\mathbb{F}$-vector space. Let $\langle,\rangle\colon V \times V \to \mathbb{F}$ be a sesquilinear form. Let $k \in \mathbb{Z}_{>0}$ and let W be a subspace of V such that $\dim(W) = k$ and $W \cap W^\perp = 0$. The orthogonal projection onto W is the unique linear transformation $P\colon V \to V$ such that*

(1) If $v \in V$ then $P(v) \in W$.

(2) If $v \in V$ and $w \in W$ then $\langle v, w\rangle = \langle P(v), w\rangle$.

Proof. Let $(w_1, \ldots, w_k)$ be a basis of W and let $(w^1, \ldots, w^k)$ be the dual basis of W. The *orthogonal projection onto W* is the function

$$P_W\colon V \to V \qquad \text{given by} \qquad P_W(v) = \sum_{i=1}^{k} \langle v, w_i\rangle w^i.$$

To show: (a) P_W is a linear transformation that satisfies conditions (1) and (2).
(b) If Q is a linear transformation that satisfies (1) and (2) then $Q = P_W$.

(a) To show: (0) P_W is a linear transformation.
(1) If $v \in V$ then $P(v) \in W$.
(2) If $v \in V$ and $w \in W$ then $\langle v, w\rangle = \langle P(v), w\rangle$.

(0) To show: If $c \in \mathbb{F}$ and $v, v_1, v_2 \in V$ then $P_W(cv) = cP_W(v)$ and $P_W(v_1 + v_2) = P_W(v_1) + P_W(v_2)$.
Assume $c \in \mathbb{F}$ and $v, v_1, v_2 \in V$.
To show: $P_W(cv) = cP_W(v)$ and $P_W(v_1 + v_2) = P_W(v_1) + P_W(v_2)$.
Since $\langle,\rangle$ ls linear in the first coordinate then

$$P_W(cv) = \sum_{i=1}^{k} \langle cv, w_i\rangle w^i = \sum_{i=1}^{k} c\langle v, w_i\rangle w^i = c\Big(\sum_{i=1}^{k} \langle v, w_i\rangle w^i\Big) = cP_W(v), \quad \text{and}$$

$$P_W(cv) = \sum_{i=1}^{k} \langle v_1 + v_2, w_i\rangle w^i = \sum_{i=1}^{k} c\langle v_1, w_i\rangle w^i + \sum_{i=1}^{k} c\langle v_1, w_i\rangle w^i = PW(v_1) + P_W(v_2).$$

So P_W is a linear transformation.

(1) Assume $v \in V$.

Since $w^1, \ldots, w^k \in W$ and $P_W(v) = \sum_{i=1}^k \langle v, w_i \rangle w^i$ then $P_W(v) \in W$.

(2) Assume $v \in V$ and $w \in W$.
Since $\{w_1, \ldots, w_k\}$ is a basis of W then there exist $c_1, \ldots, c_k \in \mathbb{F}$ such that $w = c_1 w_1 + \cdots + c_k w_k$.
Then

$$\langle P_W(v), w \rangle = \left\langle \sum_{i=1}^k \langle v, w_i \rangle w^i, \sum_{j=1}^k c_j w_j \right\rangle = \sum_{i=1}^k \overline{c_i} \langle v, w_i \rangle = \langle v, w \rangle.$$

Thus $P_W(v)$ is a linear transformation that satisfies (1) and (2).

(b) Assume $Q\colon V \to V$ is a linear transformation that satisfies (1) and (2).
To show: $Q = P_W$.

To show: If $v \in V$ then $Q(v) = P_W(v)$.
Assume $v \in V$.
Since Q satisfies property (2), if $w \in W$ then $\langle Q(v), w \rangle = \langle v, w \rangle$.
So $\langle Q(v), w \rangle = \langle v, w \rangle = \langle P_W(v), w \rangle$.
So, if $w \in W$ then $\langle P_W(v) - Q(v), w \rangle = 0$.
So $P_W(v) - Q(v) \in W^\perp$.
By Property (1), $P_W(v) - Q(v) \in W$.
So $P_W(v) - Q(v) \in W \cap W^\perp$.
Since $W \cap W^\perp = 0$ then $P_W(v) - Q(v) = 0$.

So $P_W = Q$.

□

10.12.6 Orthogonal decompositions

Theorem 10.18. *Let $n \in \mathbb{Z}_{>0}$ and let V be an inner product space with $\dim(V) = n$. Let W be a subspace of V such that $W \cap W^\perp = 0$. Let P_W be the orthogonal projection onto W and let $P_{W^\perp} = 1 - P_W$. Then*

$$P_W^2 = P_W, \quad P_{W^\perp}^2 = P_{W^\perp}, \quad P_W P_{W^\perp} = P_{W^\perp} P_W = 0, \quad 1 = P_W + P_{W^\perp},$$

$$\ker(P_W) = W^\perp, \qquad \operatorname{im}(P_W) = W \quad \text{and} \quad V = W \oplus W^\perp.$$

Proof. (a) Assume $v \in V$. Then, by properties (1) and (2) of Proposition 16.6,

$$P_W^2(v) = \sum_{i=1}^k \langle P_W(v), w^i \rangle w_i = \sum_{i=1}^k \langle v, w^i \rangle w_i = P_W(v). \qquad \text{So } P_W^2 = P_W.$$

(b) Since $P_W^2 = P_W$ then

$$P_{W^\perp}^2 = (1 - P_W)^2 = 1 - 2P_W + P_W^2 = 1 - 2P_W + P_W = 1 - P_W = P_{W^\perp}.$$

(c) Since $P_W^2 = P_W$ and $P_{W^\perp} = 1 - P_W$ then

$$P_W P_{W^\perp} = P_W(1-P_W) = P_W - P_W^2 = P_W - P_W = 0 \quad \text{and}$$
$$P_{W^\perp} P_W = (1-P_W)P_W = P_W - P_W^2 = P_W - P_W = 0.$$

(d) Since $P_{W^\perp} = 1 - P_W$ then $P_W + P_{W^\perp} = P_W + (1-P_W) = 1$.

(e) To show $\ker(P_W) = W^\perp$.

To show: (ea) $\ker(P_W) \subseteq W^\perp$.

(eb) $W^\perp \subseteq \ker(P_W)$.

(ea) Assume $v \in \ker(P_W)$.
By property (2) in Proposition 16.6, $\langle v, w\rangle = \langle P_W(v), w\rangle = \langle 0, w\rangle = 0$.
So $v \in W^\perp$.
So $\ker(P_W) \subseteq W^\perp$.

(eb) Assume $v \in W^\perp$.
If $w \in W$ then $\langle P_W(v), w\rangle = \langle v, w\rangle = 0$ and so $P_W(v) \in W^\perp$.
By property (1), $P_W(v) \in W$ and so $P_W(v) \in W \cap W^\perp = 0$.
So $v \in \ker(P_W)$.
So $W^\perp \subseteq \ker(P_W)$.

So $\ker(P_W) = W^\perp$.

(f) To show: $\mathrm{im}(P_W) = W$.

To show: (fa) $\mathrm{im}(P_W) \subseteq W$.

(fb) $W \subseteq \mathrm{im}(P_W)$.

(fa) By property (1) of Proposition 16.6, $\mathrm{im}(P_W) \subseteq W$.

(fb) Assume $w \in W$.
Let $c_1, \ldots, c_k \in \mathbb{F}$ such that $w = c_1 w^1 + \cdots + c_k w^k$.
Since $\langle w^i, w_j\rangle = \delta_{ij}$ then

$$P_W(w) = \sum_{i=1}^k \langle w, w_i\rangle w^i = \sum_{i=1}^k \sum_{j=1}^k \langle c_j w^j, w_i\rangle w^i = \sum_{j=1}^k c_j w^i = w.$$

So $W \subseteq \mathrm{im}(P_W)$.

So $\mathrm{im}(P_W) = W$.

(g) If $v \in V$ then $v = P_W(v) + (1-P_W)(v) \in W + W^\perp$.
So $V = W + W^\perp$.
By assumption $W \cap W^\perp = 0$, and so $V = W \oplus W^\perp$.

□

10.12.7 Orthonormal sequences are linearly independent

Proposition 10.19. *Let V be an $\mathbb{F}$-vector space with a Hermitian form. An orthonormal sequence $(a_1, a_2, \ldots)$ in V is linearly independent.*

Proof. Let $(a_1, a_2, \ldots)$ be an orthonormal sequence in V.
To show: $\{a_1, a_2, \ldots\}$ is linearly independent.
To show: If $\ell \in \mathbb{Z}_{>0}$ and $\mu_1 a_1 + \mu_2 a_2 + \cdots + \mu_\ell a_\ell = 0$ then $\mu_j = 0$ for $j \in \{1, 2, \ldots, \ell\}$.
Assume $\ell \in \mathbb{Z}_{>0}$ and $\mu_1 a_1 + \mu_2 a_2 + \cdots + \mu_\ell a_\ell = 0$.
To show: If $j \in \{1, \ldots, \ell\}$ then $\mu_j = 0$.
Assume $j \in \{1, \ldots, \ell\}$.
Then $0 = \langle \mu_1 a_1 + \mu_2 a_2 + \cdots + \mu_\ell a_\ell, a_j\rangle = \mu_j \langle a_j, a_j\rangle = \mu_j$.
So $\{a_1, a_2, \ldots\}$ is linearly independent. □

10.12.8 Gram-Schmidt

Theorem 10.20. *(Gram-Schmidt) Let V be an $\mathbb{F}$-vector space with a sesquilinear form $\langle,\rangle\colon V \times V \to \mathbb{F}$. Assume that $\langle,\rangle$ is nonisotropic and that $\langle,\rangle$ is Hermitian i.e.,*

(1) (Nonisotropy condition) If $v \in V$ and $\langle v, v\rangle = 0$ then $v = 0$, and

(2) (Hermitian condition) If $v_1, v_2 \in V$ then $\langle v_2, v_1\rangle = \overline{\langle v_1, v_2\rangle}$.

Let $p_1, p_2, \ldots$ be a sequence of linear independent elements of V.

(a) Define $b_1 = p_1$ and

$$b_{n+1} = p_{n+1} - \frac{\langle p_{n+1}, b_1\rangle}{\langle b_1, b_1\rangle} b_1 - \cdots - \frac{\langle p_{n+1}, b_n\rangle}{\langle b_n, b_n\rangle} b_n, \qquad \text{for } n \in \mathbb{Z}_{>0}.$$

Then $(b_1, b_2, \ldots)$ is an orthogonal sequence in V.

(b) Assume that $\mathbb{F}$ is a field in which square roots can be made sense of and that if $v \in V$ and $v \neq 0$ then $\langle v, v\rangle \neq 0$. Define

$$\|v\| = \sqrt{\langle v, v\rangle}, \qquad \text{for } v \in V.$$

Let $(b_1, \ldots, b_n)$ be an orthogonal basis of V. Define

$$u_1 = \frac{b_1}{\|b_1\|}, \quad \ldots\ , \quad u_n = \frac{b_n}{\|b_n\|}.$$

Then $(u_1, \ldots, u_n)$ is an orthonormal basis of V.

Proof. (Sketch) The proof is by induction on n.
For the base case, there is only one vector b_1 and so there is nothing to show.
Induction step: Assume $(b_1, \ldots, b_n)$ are orthogonal.
Let $j \in \{1, \ldots, n\}$. Then

$$\begin{aligned}
\langle b_{n+1}, b_j\rangle &= \left\langle p_{n+1} - \frac{\langle p_{n+1}, b_1\rangle}{\langle b_1, b_1\rangle} b_1 - \cdots - \frac{\langle p_{n+1}, b_n\rangle}{\langle b_n, b_n\rangle} b_n, b_j \right\rangle \\
&= \langle p_{n+1}, b_j\rangle - \frac{\langle p_{n+1}, b_1\rangle}{\langle b_1, b_1\rangle}\langle b_1, b_j\rangle - \cdots - \frac{\langle p_{n+1}, b_n\rangle}{\langle b_n, b_n\rangle}\langle b_n, b_j\rangle \\
&= \langle p_{n+1}, b_j\rangle - \frac{\langle p_{n+1}, b_j\rangle}{\langle b_j, b_j\rangle}\langle b_j, b_j\rangle = \langle p_{n+1}, b_j\rangle - \langle p_{n+1}, b_j\rangle = 0 \quad \text{and} \\
\langle b_j, b_{n+1}\rangle &= \left\langle b_j, p_{n+1} - \frac{\langle p_{n+1}, b_1\rangle}{\langle b_1, b_1\rangle} b_1 - \cdots - \frac{\langle p_{n+1}, b_n\rangle}{\langle b_n, b_n\rangle} b_n \right\rangle \\
&= \langle b_j, p_{n+1}\rangle - \frac{\overline{\langle p_{n+1}, b_1\rangle}}{\langle b_1, b_1\rangle}\langle b_j, b_1\rangle - \cdots - \frac{\overline{\langle p_{n+1}, b_n\rangle}}{\langle b_n, b_n\rangle}\langle b_j, b_n\rangle \\
&= \langle b_j, p_{n+1}\rangle - \frac{\overline{\langle p_{n+1}, b_j\rangle}}{\langle b_j, b_j\rangle}\langle b_j, b_j\rangle = \langle b_j, p_{n+1}\rangle - \overline{\langle p_{n+1}, b_j\rangle} = 0,
\end{aligned}$$

where the identity $\overline{\langle b_k, b_k\rangle} = \langle b_k, b_k\rangle$ and the last equality follow from the assumption that $\langle , \rangle$ is Hermitian. So $(b_1, \ldots, b_{n+1})$ are orthogonal. □

10.12.9 Normal matrices give invariance

Proposition 10.21. *Let $V = \mathbb{C}^n$ with inner product given by (17.1). Let*

$$A \in M_n(\mathbb{C}), \qquad \lambda \in \mathbb{C} \qquad \textit{and} \qquad V_\lambda = \ker(\lambda - A).$$

If $AA^* = A^*A$ then

$$V_\lambda \textit{ is } A\textit{-invariant}, \quad V_\lambda^\perp \textit{ is } A\textit{-invariant}, \quad V_\lambda \textit{ is } A^*\textit{-invariant and} \quad V_\lambda^\perp \textit{ is } A^*\textit{-invariant.}$$

Proof.

(a) Let $p \in V_\lambda$. Then $Ap = \lambda p \in V_\lambda$. So V_λ is A invariant.
(b) Let $p \in V_\lambda$. Since $A(A^*p) = A^*Ap = \lambda A^*p$ then $A^*p \in V_\lambda$. So V_λ is A^* invariant.
(c) Let $z \in V_\lambda^\perp$.
To show $Az_\lambda \in V_\lambda^\perp$.
To show: If $u \in V_\lambda$ then $\langle Az, u\rangle = 0$.
Assume $u \in V_\lambda$.
To show: $\langle Az, u\rangle = 0$.
By (b), $A^*u \in V_\lambda$, and so $\langle Az, u\rangle = \langle z, A^*u\rangle = 0$.
So $Az \in V_\lambda^\perp$.
So $V_\lambda^\perp$ is A-invariant.
(d) Let $z \in V_\lambda^\perp$.
To show: If $u \in V_\lambda$ then $\langle A^*z, u\rangle = 0$.

$$\langle A^*z, u\rangle = \langle z, Au\rangle = 0, \qquad \text{since } Au \in V_\lambda.$$

So $A^*z \in V_\lambda^\perp$. So $V_\lambda^\perp$ is A^*-invariant.

□

10.12.10 The spectral theorem

Theorem 10.22. *(Spectral theorem)*
Let $n \in \mathbb{Z}_{>0}$ and $V = \mathbb{C}^n$ with inner product given by (17.1).

(a) Let $n \in \mathbb{Z}_{>0}$ and $A \in M_n(\mathbb{C})$ such that $AA^* = A^*A$. Then there exists a unitary $U \in M_n(\mathbb{C})$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ such that

$$U^{-1}AU = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

(b) Let $f\colon V \to V$ be a linear transformation such that $ff^* = f^*f$. Then there exists an orthonormal basis $(u_1, \ldots, u_n)$ of V consisting of eigenvectors of f.

Proof. The two statements are equivalent via the relation between A and f given by

$$\begin{array}{rccc} f\colon & V & \longrightarrow & V \\ & v & \longmapsto & Av. \end{array}$$

The proof is by induction on n.
The base case is when $\dim(V) = 1$. In this case $A \in M_1(\mathbb{C})$ is diagonal.
The induction step:
For $\mu \in \mathbb{C}$ let $V_\mu = \ker(\mu - f)$, the μ-eigenspace of f.
Since $\mathbb{C}$ is algebraically closed, there exists $\lambda \in \mathbb{C}$ which is a root of the characteristic polynomial $\det(x - A)$.
So there exists $\lambda \in \mathbb{C}$ such that $\det(\lambda - A) = 0$.
So there exists $\lambda \in \mathbb{C}$ such that $V_\lambda = \ker(\lambda - A) \neq 0$.
Let $k = \dim(V_\lambda)$ and let $(p_1, \ldots, p_k)$ be a basis of V_λ.
Use Gram-Schmidt to convert $(p_1, \ldots, p_k)$ to an orthogonal basis $(u_1, \ldots, u_k)$ of V_λ.
By definition of V_λ, the basis vectors $(u_1, \ldots, u_k)$ are all eigenvectors of f (of eigenvalue λ.
By Theorem 10.18 (orthogonal decomposition) and Proposition 10.21,

$$V = V_\lambda \oplus (V_\lambda)^\perp \quad \text{and } V_\lambda^\perp \text{ is } A\text{-invariant and } A^*\text{-invariant.}$$

Let

$$f_1 \colon \begin{array}{ccc} V_\lambda^\perp & \to & V_\lambda^\perp \\ v & \mapsto & Av \end{array} \qquad \text{and} \qquad g_1 \colon \begin{array}{ccc} V_\lambda^\perp & \to & V_\lambda^\perp \\ v & \mapsto & A^*v \end{array}$$

Then $g_1 = f_1^*$ and $f_1 f_1^* = f_1^* f_1$.
Thus, by induction, there exists an orthonormal basis $(u_{k+1}, \ldots, u_n)$ of $V_\lambda^\perp$ consisting of eigenvectors of f_1.
By definition of f_1, eigenvectors of f_1 are eigenvectors of f.
So $(u_1, \ldots, u_k, u_{k+1}, \ldots, u_n)$ is an orthonormal basis of eigenvectors of f. $\square$