Chapter 7

Operators on Inner Product Spaces

7.A Self-Adjoints and Normal Operators

10/7: • **Adjoint** (of $T \in \mathcal{L}(V,W)$): The function $T^* : W \to V$ that satisfies

$$\langle Tv, w\rangle = \langle v, T^*w\rangle$$

for all $v \in V$ and $w \in W$[1].

  – Calculating T^*w: Consider the linear functional $\varphi : V \to \mathbb{F}$ defined by $\varphi(v) = \langle Tv, w\rangle$ for all $v \in V$. By the Riesz Representation Theorem, there exists a unique vector $T^*w \in V$ such that $\varphi(v) = \langle v, T^*w\rangle$ for all $v \in V$. This vector in V will guarantee that $\langle Tv, w\rangle = \varphi(v) = \langle v, T^*w\rangle$ for all $v \in V$, and we can find vectors $T^*w \in V$ for all $w \in W$.

• The adjoint is a linear map.

Theorem 7.1. *If $T \in \mathcal{L}(V,W)$, then $T^* \in \mathcal{L}(W,V)$.*

Proof. Let $T \in \mathcal{L}(V,W)$, let $w_1, w_2 \in W$, and let $\lambda \in \mathbb{F}$. By the definition of T^*, we have that for any $v \in V$,

$$\begin{aligned}\langle v, T^*(w_1+w_2)\rangle &= \langle Tv, w_1+w_2\rangle \\ &= \langle Tv, w_1\rangle + \langle Tv, w_2\rangle \\ &= \langle v, T^*w_1\rangle + \langle v, T^*w_2\rangle \\ &= \langle v, T^*w_1 + T^*w_2\rangle\end{aligned} \qquad \begin{aligned}\langle v, T^*(\lambda w_1)\rangle &= \langle Tv, \lambda w_1\rangle \\ &= \bar{\lambda}\langle Tv, w_1\rangle \\ &= \bar{\lambda}\langle v, T^*w_1\rangle \\ &= \langle v, \lambda T^*w_1\rangle\end{aligned}$$

Thus, by the definition of T^*,

$$T^*(w_1+w_2) = T^*w_1 + T^*w_2 \qquad T^*(\lambda w_1) = \lambda T^*w$$

so T^* is a linear map, as desired. ■

• Properties of the adjoint.

Theorem 7.2.

(a) $(S+T)^* = S^* + T^*$ *for all* $S < T \in \mathcal{L}(V,W)$.

[1] Note that the word adjoint has another, unrelated meaning in algebra. Fortunately, this other meaning will not be covered in **bib:Axler**.

Proof. Suppose $S, T \in \mathcal{L}(V, W)$. If $v \in V$ and $w \in W$, then

$$\begin{aligned}\langle v, (S+T)^* w\rangle &= \langle (S+T)v, w\rangle \\ &= \langle Sv, w\rangle + \langle Tv, w\rangle \\ &= \langle v, S^* w\rangle + \langle v, T^* w\rangle \\ &= \langle v, S^* w + T^* w\rangle\end{aligned}$$

Thus, $(S+T)^* w = S^* w + T^* w$, as desired. ■

(*b*) $(\lambda T)^* = \bar{\lambda} T^*$ *for all* $\lambda \in \mathbb{F}$ *and* $T \in \mathcal{L}(V, W)$.

Proof. Suppose $T \in \mathcal{L}(V, W)$ and $\lambda \in \mathbb{F}$. If $v \in V$ and $w \in W$, then

$$\begin{aligned}\langle v, (\lambda T)^* w\rangle &= \langle \lambda T v, w\rangle \\ &= \lambda \langle Tv, w\rangle \\ &= \lambda \langle v, T^* w\rangle \\ &= \langle v, \bar{\lambda} T^* w\rangle\end{aligned}$$

Thus, $(\lambda T)^* w = \bar{\lambda} T^* w$, as desired. ■

(*c*) $(T^*)^* = T$ *for all* $T \in \mathcal{L}(V, W)$.

Proof. Suppose $T \in \mathcal{L}(V, W)$. If $v \in V$ and $w \in W$, then

$$\begin{aligned}\langle w, (T^*)^* v\rangle &= \langle T^* w, v\rangle \\ &= \overline{\langle v, T^* w\rangle} \\ &= \overline{\langle Tv, w\rangle} \\ &= \langle w, Tv\rangle\end{aligned}$$

Thus, $(T^*)^* v = Tv$, as desired. ■

(*d*) $I^* = I$, *where* I *is the identity operator on* V.

Proof. If $v, u \in V$, then

$$\langle v, I^* u\rangle = \langle Iv, u\rangle = \langle v, Iu\rangle$$

Thus, $I^* u = Iu$, as desired. ■

(*e*) $(ST)^* = T^* S^*$ *for all* $T \in \mathcal{L}(V, W)$ *and* $S \in \mathcal{L}(W, U)$. *Here* U *is an inner product space over* $\mathbb{F}$.

Proof. Suppose $T \in \mathcal{L}(V, W)$ and $S \in \mathcal{L}(W, U)$. If $v \in V$ and $u \in U$, then

$$\begin{aligned}\langle v, (ST)^* u\rangle &= \langle STv, u\rangle \\ &= \langle Tv, S^* u\rangle \\ &= \langle v, T^* S^* u\rangle\end{aligned}$$

Thus, $(ST)^* u = T^* S^* u$, as desired. ■

- Null space and range of T^*.

Theorem 7.3. *Suppose* $T \in \mathcal{L}(V, W)$. *Then*

(*a*) $\operatorname{null} T^* = (\operatorname{range} T)^\perp$.

Proof. Let $w \in W$ be an arbitrary element of $\operatorname{null} T^*$. Then $T^* w = 0$ by definition. It follows by Theorem 6.1c that $\langle v, T^* w\rangle = 0$ for all $v \in V$. Thus, by the definition of the adjoint, $\langle Tv, w\rangle = 0$ for all $v \in V$. But this implies that w is orthogonal to every vector in $\operatorname{range} T$ (i.e., the set of all Tv), meaning that $w \in (\operatorname{range} T)^\perp$.
The proof is symmetric in the other direction. ■

(b) $\operatorname{range} T^* = (\operatorname{null} T)^\perp$.

Proof. We have that

$$\begin{aligned} \operatorname{range} T^* &= ((\operatorname{range} T^*)^\perp)^\perp && \text{Theorem 6.22} \\ &= (\operatorname{null}(T^*)^*)^\perp && \text{Theorem 7.3a} \\ &= (\operatorname{null} T)^\perp && \text{Theorem 7.2c} \end{aligned}$$

as desired. ■

(c) $\operatorname{null} T = (\operatorname{range} T^*)^\perp$.

Proof. We have that

$$\begin{aligned} \operatorname{null} T &= \operatorname{null}(T^*)^* && \text{Theorem 7.2c} \\ &= (\operatorname{range} T^*)^\perp && \text{Theorem 7.3a} \end{aligned}$$

as desired. ■

(d) $\operatorname{range} T = (\operatorname{null} T^*)^\perp$.

Proof. We have that

$$\begin{aligned} \operatorname{range} T &= ((\operatorname{range} T)^\perp)^\perp && \text{Theorem 6.22} \\ &= (\operatorname{null} T^*)^\perp && \text{Theorem 7.3a} \end{aligned}$$

as desired. ■

- **Conjugate transpose** (of an m-by-n matrix): The n-by-m matrix obtained by interchanging the rows and columns and then taking the complex conjugate of each entry.
  - "If $\mathbb{F} = \mathbb{R}$, then the conjugate transpose of a matrix is the same as its transpose" (**bib:Axler**).
- The next result shows how to compute the matrix of T^* from the matrix of T. Note, however, that if $\mathcal{M}(T)$ is with respect to nonorthonormal bases, $\mathcal{M}(T^*)$ does not necessarily equal the conjugate transpose of $\mathcal{M}(T)$.

Theorem 7.4. *Let $T \in \mathcal{L}(V, W)$. Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $f_1, \dots, f_m$ is an orthonormal basis of W. Then*

$$\mathcal{M}(T^*, (f_1, \dots, f_m), (e_1, \dots, e_n))$$

is the conjugate transpose of

$$\mathcal{M}(T, (e_1, \dots, e_n), (f_1, \dots, f_m))$$

Proof. Recall that the k^{th} column of $\mathcal{M}(T)$ is given by writing Te_k as a linear combination of the f_j's. Since $f_1, \dots, f_m$ is an orthonormal basis of W, Theorem 6.12 implies that

$$Te_k = \langle Te_k, f_1 \rangle f_1 + \cdots + \langle Te_k, f_m \rangle f_m$$

Thus, the entry in row j column k of $\mathcal{M}(T)$ is $\langle Te_k, f_j \rangle$. On the other hand, since

$$T^* f_k = \langle T^* f_k, e_1 \rangle e_1 + \cdots + \langle T^* f_k, e_n \rangle e_n$$

we have that the entry in row j column k of $\mathcal{M}(T^*)$ is

$$\begin{aligned} \langle T^* f_k, e_j \rangle &= \langle f_k, Te_j \rangle \\ &= \overline{\langle Te_j, f_k \rangle} \end{aligned}$$

Therefore, the entry in row k column j of $\mathcal{M}(T^*)$ is the complex conjugate of the entry in row j column k of $\mathcal{M}(T)$, as desired. ■

- **Self-adjoint** (operator $T \in \mathcal{L}(V)$): An operator T such that $T = T^*$. *Also known as* **Hermitian**.
  - In other words, $T \in \mathcal{L}(V)$ is self-adjoint if and only if

    $$\langle Tv, w \rangle = \langle v, Tw \rangle$$

    for all $v, w \in V$.
- The sum of two self-adjoint operators is self-adjoint, and the product of a real scalar and a self-adjoint operator is self-adjoint.
- Note the analogy between self-adjoint operator and complex numbers: A complex number z is real iff $z = \bar{z}$, and thus a self-adjoint operator ($T = T^*$) is analogous to a real number.
- Eigenvalues of self-adjoint operators.

  Theorem 7.5. *Every eigenvalue of a self-adjoint operator is real.*

  Proof. Let T be a self-adjoint operator on V, let λ be an eigenvalue of T, and let v be a nonzero vector in V such that $Tv = \lambda v$. Then

  $$\lambda \|v\|^2 = \langle \lambda v, v \rangle = \langle Tv, v \rangle = \langle v, Tv \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \|v\|^2$$

  so $\lambda = \bar{\lambda}$, which implies that λ is real, as desired. ∎

- The next result is false for real inner product spaces (consider a rotation matrix), but true for complex ones.

  Theorem 7.6. *Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Suppose $\langle Tv, v \rangle = 0$ for all $v \in V$. Then $T = 0$.*

  Proof. Let $u \in V$ be arbitrary. By inner product algebra, we have that

  $$\langle Tu, w \rangle = \frac{\langle T(u+w), u+w \rangle - \langle T(u-w), u-w \rangle}{4} + \frac{\langle T(u+iw), u+iw \rangle - \langle T(u-iw), u-iw \rangle}{4} i$$

  for all $w \in V$. Since each term on the right-hand side of the above equation is of the form $\langle Tv, v \rangle$ and we know by hypothesis that $\langle Tv, v \rangle = 0$ for all $v \in V$, we have that $\langle Tu, w \rangle = 0$ for all $w \in V$. In particular, if we let $w = Tu$, we learn that $\langle Tu, Tu \rangle = 0$, which implies that $Tu = 0$. But this implies that $Tu = 0$ for all $u \in V$, i.e., that $T = 0$. ∎

- The next result provides another example of how self-adjoint operators behave like real numbers, and is also false for real inner product spaces (consider a operator on such a space that is not self-adjoint).

  Theorem 7.7. *Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$. Then T is self-adjoint if and only if $\langle Tv, v \rangle \in \mathbb{R}$ for every $v \in V$.*

  Proof. Suppose first that T is self-adjoint. Let $v \in V$ be arbitrary. Then

  $$\langle Tv, v \rangle - \overline{\langle Tv, v \rangle} = \langle Tv, v \rangle - \langle v, Tv \rangle = \langle Tv, v \rangle - \langle T^* v, v \rangle = \langle (T - T^*)v, v \rangle = \langle 0v, v \rangle = 0$$

  so $\langle Tv, v \rangle = \overline{\langle Tv, v \rangle}$. Therefore, $\langle Tv, v \rangle \in \mathbb{R}$, as desired.

  Now suppose that $\langle Tv, v \rangle \in \mathbb{R}$ for every $v \in V$. Let $v \in V$ be arbitrary. Then

  $$\langle (T - T^*)v, v \rangle = \langle Tv, v \rangle - \langle T^* v, v \rangle = \langle Tv, v \rangle - \langle v, Tv \rangle = \langle Tv, v \rangle - \overline{\langle Tv, v \rangle} = 0$$

  Therefore, by Theorem 7.6, $T - T^* = 0$, or $T = T^*$, as desired. ∎

- We now show that on complex *or* real vector spaces, self-adjoint operators that satisfy $\langle Tv, v\rangle = 0$ *must* be the zero operator.

  Theorem 7.8. *Suppose T is a self-adjoint operator on V such that*
  $$\langle Tv, v\rangle = 0$$
  for all $v \in V$. Then $T = 0$.

  Proof. We divide into two cases. If V is complex, invoke Theorem 7.6. If V is real, we continue.

  Let $u \in V$ be arbitrary. By inner product algebra, we have that
  $$\langle Tu, w\rangle = \frac{\langle T(u+w), u+w\rangle - \langle T(u-w), u-w\rangle}{4}$$
  By a symmetric argument to that used in the later part of the proof of Theorem 7.6, we can confirm that $T = 0$. ∎

- **Normal** (operator): An operator that commutes with its adjoint.
  - In other words, $T \in \mathcal{L}(V)$ is normal if
  $$TT^* = T^*T$$
- Every self-adjoint operator is normal.
- We now characterize normal operators.

  Theorem 7.9. *An operator is normal if and only if*
  $$\|Tv\| = \|T^*v\|$$
  for all $v \in V$.

  Proof. Let $T \in \mathcal{L}(V)$.

  Suppose first that T is normal. Then $T^*T - TT^* = 0$. Thus, by Theorem 6.1b, $\langle (T^*T - TT^*)v, v\rangle = 0$ for all $v \in V$. It follows that
  $$\begin{aligned}
  \langle T^*Tv, v\rangle &= \langle TT^*v, v\rangle \\
  \langle Tv, Tv\rangle &= \langle T^*v, T^*v\rangle \\
  \|Tv\|^2 &= \|T^*v\|^2 \\
  \|Tv\| &= \|T^*v\|
  \end{aligned}$$
  for all $v \in V$, as desired.

  Now suppose that $\|Tv\| = \|T^*v\|$ for all $v \in V$. Then following the reverse of the procedure for the forward direction, we can easily show that $\langle (T^*T - TT^*)v, v\rangle = 0$ for all $v \in V$. Additionally, by consecutive applications of Theorem 7.2, we have that
  $$\begin{aligned}
  (T^*T - TT^*)^* &= (T^*T)^* - (TT^*)^* \\
  &= T^*(T^*)^* - (T^*)^*T^* \\
  &= T^*T - TT^*
  \end{aligned}$$
  It follows that $T^*T - TT^*$ is self-adjoint. This combined with the previous result implies by Theorem 7.8 that $T^*T - TT^* = 0$. It follows that $T^*T = TT^*$, so T is normal, as desired. ∎

- While an operator and its adjoint may have different eigenvectors, a normal operator and its adjoint have the same eigenvectors.

Theorem 7.10. *Suppose $T \in \mathcal{L}(V)$ is normal and $v \in V$ is an eigenvector of T with eigenvalue λ. Then v is also an eigenvector of T^* with eigenvalue $\bar{\lambda}$.*

Proof. By consecutive applications of Theorem 7.2, we have that

$$\begin{aligned}(T-\lambda I)(T-\lambda I)^* &= (T-\lambda I)(T^*-\bar{\lambda}I)\\ &= TT^*-\bar{\lambda}T-\lambda T^*+\lambda\bar{\lambda}I\\ &= T^*T-\lambda T^*-\bar{\lambda}T+\bar{\lambda}\lambda I\\ &= (T^*-\bar{\lambda}I)(T-\lambda I)\\ &= (T-\lambda I)^*(T-\lambda I)\end{aligned}$$

Thus, $T-\lambda I$ is self-adjoint. It follows by Theorem 7.9 that

$$0 = \|(T-\lambda I)v\| = \|(T-\lambda I)^*v\| = \left\|(T^*-\bar{\lambda}I)v\right\|$$

Hence $(T^*-\bar{\lambda}I)v = 0$, so $T^*v = \bar{\lambda}v$, so v is an eigenvector of T^* with eigenvalue $\bar{\lambda}$, as desired. ■

- Normal operators have orthogonal eigenvectors.

Theorem 7.11. *Suppose $T \in \mathcal{L}(V)$ is normal. Then the eigenvectors of T corresponding to distinct eigenvalues are orthogonal.*

Proof. Let α, β be distinct eigenvalues of T, and let u, v be their corresponding eigenvectors. Thus, we have that

$$\begin{aligned}(\alpha-\beta)\langle u,v\rangle &= \langle \alpha u, v\rangle - \langle u, \bar{\beta}v\rangle\\ &= \langle Tu, v\rangle - \langle u, T^*v\rangle && \text{Theorem 7.10}\\ &= 0\end{aligned}$$

Since $\alpha \neq \beta$ by hypothesis, we must have that $\langle u,v\rangle = 0$. Therefore, u, v are orthogonal, as desired. ■

7.B The Spectral Theorem

- Diagonal operators are nice operators.
  - An operator has a diagonal matrix with respect to some basis iff the basis consists of eigenvectors of the operator (see Theorem 5.11).
- The nicest operators are those for which there is an orthonormal basis of V with respect to which the operator has a diagonal matrix.
  - The Spectral Theorem characterizes the operators $T \in \mathcal{L}(V)$ for which there exists an orthonormal basis of V consisting of eigenvectors of T.
  - In particular, it characterizes them as the normal operators when $\mathbb{F} = \mathbb{C}$ and the self-adjoint operators when $\mathbb{F} = \mathbb{R}$.
  - "The Spectral Theorem is probably the most useful tool in the study of operators on inner product spaces" (**bib:Axler**).
- For the purposes of proving the Spectral Theorem, we will break it into a Complex Spectral Theorem and a Real Spectral Theorem.
- The complex portion is simpler, so we begin with it.

Theorem 7.12 (Complex Spectral Theorem)**.** *Suppoe $\mathbb{F} = \mathbb{C}$ and $T \in \mathcal{L}(V)$. Then the following are equivalent.*

(a) T is normal.

(b) V has an orthonormal basis consisting of eigenvectors of T.

(c) T has a diagonal matrix with respect to some orthonormal basis of V.

Proof. We have by Theorem 5.11 that (b) and (c) are equivalent, so we will focus on proving the equivalence of (a) and (c).

Suppose first that (c) holds. Since $\mathcal{M}(T)$ is diagonal and $\mathcal{M}(T^*)$ is the conjugate transpose of $\mathcal{M}(T)$, $\mathcal{M}(T^*)$ is diagonal. Therefore, since any two diagonal matrices commute, T is normal, so (a) holds.

Now suppose that (a) holds. By Schur's Theorem, there exists an orthonormal basis $e_1, \dots, e_n$ of V with respect to which T has an upper triangular matrix. We will show that this matrix is actually diagonal. To begin, since $\mathcal{M}(T)$ is upper triangular, we know that

$$\|Te_1\|^2 = |a_{1,1}|^2$$

Similarly, since T^* is the conjugate *transpose*, we have that

$$\|T^*e_1\|^2 = |a_{1,1}|^2 + \cdots + |a_{1,n}|^2$$

But since $\|Te_1\| = \|T^*e_1\|$ by Theorem 7.9, the two equations above imply that

$$0 = |a_{1,2}|^2 + \cdots + |a_{1,n}|^2$$

Therefore, we know that all entries in row 1 save the first are zero. We may repeat this procedure for every row to finish the proof. ■

- The next result continues to build on the likeness of normal matrices and real numbers. Specifically, it plays off the fact that if $b, c \in \mathbb{R}$ with $b^2 < 4c$, then $x^2 + bx + c > 0$, i.e., $x^2 + bx + c$ nonzero is an "invertible" real number.

Theorem 7.13. *Suppose $T \in \mathcal{L}(V)$ is self-adjoint and $b, c \in \mathbb{R}$ are such that $b^2 < 4c$. Then*

$$T^2 + bT + cI$$

is invertible.

Proof. To prove that $T^2 + bT + cI$ is invertible, Theorem 3.18 tells us that it will suffice to show that T is injective. To do this, Theorem 3.4 tells us that we must verify that $\operatorname{null}(T^2 + bT + cI) \subset \{0\}$, i.e., that if $v \in V$ is nonzero, then $(T^2 + bT + cI)v \neq 0$. Let's begin.

Let $v \in V$ be arbitrary. Then we have that

$$\begin{aligned}
\left\langle (T^2 + bT + cI)v, v \right\rangle &= \left\langle T^2 v, v \right\rangle + b\left\langle Tv, v \right\rangle + c\left\langle v, v \right\rangle \\
&= \left\langle Tv, Tv \right\rangle + b\left\langle Tv, v \right\rangle + c\|v\|^2 \\
&\geq \|Tv\|^2 - |b|\|Tv\|\|v\| + c\|v\|^2 && \text{Cauchy-Schwarz Inequality} \\
&= \left(\|Tv\| - \frac{|b|\|v\|}{2}\right)^2 + \left(c - \frac{b^2}{4}\right)\|v\|^2 \\
&> 0
\end{aligned}$$

The overall strict inequality implies by the contrapositive of Theorem 6.1b that $(T^2 + bT + cI)v \neq 0$, as desired. ■

- Like Theorem 5.5 told us that operators on *finite-dimensional nonzero complex* vector spaces have eigenvalues, the following tells us that *self-adjoint* operators on *any nonzero* vector space have eigenvalues.

Theorem 7.14. *Suppose $V \neq \{0\}$ and $T \in \mathcal{L}(V)$ is a self-adjoint operator. Then T has an eigenvalue.*

Proof. Let V be a real inner product space, let $n = \dim V$, and let $v \in V$ be arbitrary and nonzero. Since $v, Tv, T^2v, \dots, T^nv$ has length $n+1 > \dim V$, it is linearly dependent. Thus, there exist $a_0, \dots, a_n \in \mathbb{F}$ such that

$$0 = a_0v + a_1Tv + \cdots + a_nT^nv$$

If we let the a's be the coefficients of a degree n polynomial, then we have by Theorem 4.9 that

$$a_0 + a_1x + \cdots + a_nx^n = c(x^2 + b_1x + c_1) \cdots (x^2 + b_Mx + c_M)(x - \lambda_1) \cdots (x - \lambda_m)$$

where $c \in \mathbb{R}$ is nonzero, each $b_j, c_j, \lambda_j \in \mathbb{R}$, each $b_j^2 < 4c_j$, $m + M \geq 1$, and the equation holds for all $x \in \mathbb{R}$. It follows that

$$\begin{aligned}
0 &= a_0v + a_1Tv + \cdots + a_nT^nv \\
&= (a_0I + a_1T + \cdots + a_nT^n)v \\
&= c(T^2 + b_1T + c_1I) \cdots (T_2 + b_MT + c_MI)(T - \lambda_1I) \cdots (T - \lambda_mI)v
\end{aligned}$$

Since T is self-adjoint and $b_j, c_j \in \mathbb{R}$ satisfy $b_j^2 < 4c_j$ for each j, we have by consecutive applications of Theorem 7.13 that each $T^2 + b_jT + c_jI$ is invertible. Thus, if we multiply both sides of the above equation by $1/c$ (recall that $c \neq 0$) and $(T^2 + b_jT + c_jI)^{-1}$ for each j, we obtain

$$0 = (T - \lambda I) \cdots (T - \lambda_mI)v$$

Therefore, by an argument symmetric to that used in the last paragraph of the proof of Theorem 5.5, we have that T has an eigenvalue, as desired. ■

- Invariant subspaces and self-adjoint operators.

Theorem 7.15. *Suppose $T \in \mathcal{L}(V)$ is self-adjoint and U is a subspace of V that is invariant under T. Then*

(a) *$U^\perp$ is invariant under T.*

Proof. Let $v \in U^\perp$ be arbitrary, and let u be any element of U. Then

$$\langle Tv, u \rangle = \langle v, Tu \rangle = 0$$

where the first equality holds because T is self-adjoint and the second equality holds because U is invariant under T (so $Tu \in U$, and we know that the inner product of an element of $U^\perp$ with an element of U is 0). Thus, since $\langle Tv, u \rangle = 0$ for all $u \in U$, $Tv \in U^\perp$, as desired. ■

(b) *$T|_U \in \mathcal{L}(U)$ is self-adjoint.*

Proof. If $u, v \in U$, then

$$\langle (T|_U)u, v \rangle = \langle Tu, v \rangle = \langle u, Tv \rangle = \langle u, (T|_U)v \rangle$$

as desired. ■

(c) *$T|_{U^\perp} \in \mathcal{L}(U^\perp)$ is self-adjoint.*

Proof. The proof is symmetric to that of Theorem 7.15b. ■

- We can now prove the real portion of the spectral theorem.

Theorem 7.16 (Real Spectral Theorem)**.** *Suppose $\mathbb{F} = \mathbb{R}$ and $T \in \mathcal{L}(V)$. Then the following are equivalent.*

(*a*) *T is self-adjoint.*

(*b*) *V has an orthonormal basis consisting of eigenvectors of T.*

(*c*) *T has a diagonal matrix with respect to some orthonormal basis of V.*

Proof. We will prove that (a) implies (b), (b) implies (c), and (c) implies (a). Let's begin.

First, suppose that T is self-adjoint. We induct on $\dim V$. For the base case $\dim V = 1$, we must have $Tv = \lambda v$ for any $v \in V$. Thus, take $e = v/\|v\|$ as an orthonormal basis of V consisting of eigenvectors of T. Now suppose inductively that (a) implies (b) for all real inner product spaces of dimension less than $\dim V > 1$. Suppose $T \in \mathcal{L}(V)$ is self-adjoint. By Theorem 7.14, we may let v be an eigenvector of T. It follows that $u = v/\|v\|$ is a normal eigenvector of T. Let $U = \operatorname{span}(u)$. Then U is a subspace of V that is invariant under T, so we have by Theorem 7.15c that $T|_{U^\perp} \in \mathcal{L}(U^\perp)$ is self-adjoint. But since $\dim U^\perp = \dim V - \dim U = \dim V - 1$, we have by the inductive hypothesis that there is an orthonormal basis of $U^\perp$ consisting of eigenvectors of $T|_{U^\perp}$. Adjoining u to this list gives an orthonormal basis of V consisting of eigenvectors of T, as desired.

Second, suppose that V has an orthonormal basis $e_1, \dots, e_n$ consisting of eigenvectors of T. Then since

$$Te_j = 0e_1 + \cdots + 0e_{j-1} + \lambda_j e_j + 0e_{j+1} + \cdots + 0e_n$$

for all j, we have by the definition that $\mathcal{M}(T, (e_1, \dots, e_n))$ is diagonal, as desired.

Third, suppose that T has a diagonal matrix $\mathcal{M}(T)$ with respect to some orthonormal basis of V. In a real inner product space, $\overline{\mathcal{M}(T)} = \mathcal{M}(T)$. Additionally, any diagonal matrix is equal to its transpose. Thus, $T = T^*$, so T is self-adjoint, as desired. ■

7.C Positive Operators and Isometries

10/11: • **Positive** ($T \in \mathcal{L}(V)$): A self-adjoint operator $T \in \mathcal{L}(V)$ such that

$$\langle Tv, v\rangle \geq 0$$

for all $v \in V$. *Also known as* **positive semidefinite** (operator).

  - Note that if V is complex, Theorem 7.7 implies based on the condition that $\langle Tv, v\rangle \geq 0$ for all $v \in V$ that T is self-adjoint. Therefore, in this case, we need not explicitly postulate that T is self-adjoint.

• **Square root** (of $T \in \mathcal{L}(V)$): An operator R such that $R^2 = T$.

• The following characterization of positive operators is directly analogous to the characterization of nonnegative complex numbers.

Theorem 7.17. *Let $T \in \mathcal{L}(V)$. Then the following are equivalent.*

(*a*) *T is positive.*

(*b*) *T is self-adjoint and all the eigenvalues of T are nonnegative.*

(*c*) *T has a positive square root.*

(*d*) *T has a self-adjoint square root.*

(*e*) *There exists an operator $R \in \mathcal{L}(V)$ such that $T = R^*R$.*

Proof. We will prove that (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (e) $\Rightarrow$ (a). Let's begin.

First, suppose that T is positive. Then by definition, T is self-adjoint. Additionally, let $\lambda \in \mathbb{F}$ be an eigenvalue of T. It follows by the definition of positive operators and by the positivity of the inner product that

$$\begin{aligned} 0 &\leq \langle Tv, v\rangle = \langle \lambda v, v\rangle = \lambda \langle v, v\rangle \\ 0 &\leq \lambda \end{aligned}$$

as desired.

Second, suppose that T is self-adjoint and all the eigenvalues of T are nonnegative. Since T is self-adjoint, the Real and Complex Spectral Theorems imply that there exists an orthonormal basis $e_1,\dots,e_n$ of V consisting of eigenvectors of T. Let $\lambda_1,\dots,\lambda_n$ be the corresponding eigenvalues; it follows by hypothesis that $\lambda_j \geq 0$ for all j. We now define $R \in \mathcal{L}(V)$ by

$$Re_j = \sqrt{\lambda_j}e_j$$

for all j. To prove that R is positive, let $v \in V$ be arbitrary. Suppose $v = a_1e_1 + \cdots + a_ne_n$ where $a_1,\dots,a_n \in \mathbb{F}$. Then

$$\begin{aligned}
\langle Rv, v\rangle &= \langle R(a_1e_1 + \cdots + a_ne_n), a_1e_1 + \cdots + a_ne_n\rangle \\
&= \left\langle \sqrt{\lambda_1}a_1e_1 + \cdots + \sqrt{\lambda_n}a_ne_n, a_1e_1 + \cdots + a_ne_n \right\rangle \\
&= \sum_{i=1}^n \sum_{j=1}^n \left\langle \sqrt{\lambda_i}a_ie_i, a_je_j \right\rangle \\
&= \sum_{i=1}^n \left\langle \sqrt{\lambda_i}a_ie_i, a_ie_i \right\rangle \\
&= \sum_{i=1}^n \sqrt{\lambda_i} \\
&\geq 0
\end{aligned}$$

as desired. Furthermore, $R^2e_j = \lambda_je_j = Te_j$ for each j, so by Theorem 3.1, $R^2 = T$, as desired.

Third, suppose that T has a positive square root R. Then by the definition of a positive operator, R is self-adjoint as well, as desired.

Fourth, suppose that T has a self-adjoint square root R. Since R is self-adjoint, $R = R^*$. Therefore,

$$T = R^2 = R^*R$$

as desired.

Fifth, suppose that there exists an operator $R \in \mathcal{L}(V)$ such that $T = R^*R$. To prove that T is positive, it will suffice to show that it is self-adjoint and that $\langle Tv, v\rangle \geq 0$ for all $v \in V$. First off, T is self-adjoint since

$$T^* = (R^*R)^* = R^*(R^*)^* = R^*R = T$$

Second, we have that

$$\langle Tv, v\rangle = \langle R^*Rv, v\rangle = \langle Rv, Rv\rangle \geq 0$$

for all $v \in V$. Therefore, T is positive, as desired. ■

- Since each nonnegative number has a unique nonnegative square root, the next result makes sense by analogy.

Theorem 7.18. *Every positive operator on V has a unique positive square root.*

Proof. Let T be a positive operator on V, let $v \in V$ be an eigenvector of T, let $\lambda \in \mathbb{F}$ be the corresponding eigenvalue, and let R be a positive square root of T (Theorem 7.17 guarantees that at least one such operator exists). Since T is positive, Theorem 7.17 implies that $\lambda \geq 0$. Thus, to prove that R is unique, we will prove that $Rv = \sqrt{\lambda}v$. This will imply that the behavior of R on the eigenvectors of T is uniquely determined. It will follows since there is a basis of V consisting of the eigenvectors of T (by the Real and Complex Spectral Theorems), the behavior of R on V (and hence R) is uniquely determined. Let's begin.

Since R is positive (hence self-adjoint), the Real and Complex Spectral Theorems assert that there exists an orthonormal basis $e_1, \dots, e_n$ of V consisting of eigenvectors of R. Additionally, because R is positive, the corresponding eigenvalues $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$ are nonnegative.

Now let

$$v = a_1 e_1 + \cdots + a_n e_n$$

for $a_1, \dots, a_n \in \mathbb{F}$. Then

$$a_1 \lambda e_1 + \cdots + a_n \lambda e_n = Tv = R^2 v = a_1 \lambda_1 e_1 + \cdots + a_n \lambda_n e_n$$

so since $e_1, \dots, e_n$ is linearly independent, $a_j(\lambda - \lambda_j) = 0$ for all j. It follows that

$$v = \sum_{\{j:\lambda_j = \lambda\}} a_j e_j$$

so that

$$\begin{aligned} Rv &= \sum_{j=1}^{n} a_j \sqrt{\lambda_j} e_j \\ &= \sum_{\{j:\lambda_j = \lambda\}} a_j \sqrt{\lambda_j} e_j \\ &= \sum_{\{j:\lambda_j = \lambda\}} a_j \sqrt{\lambda} e_j \\ &= \sqrt{\lambda} v \end{aligned}$$

as desired. ■

- **Isometry**: An operator $S \in \mathcal{L}(V)$ such that

$$\|Sv\| = \|v\|$$

for all $v \in V$.

  - In other words, an isometry is an operator that preserves norms.

- **Orthogonal** (operator): An isometry on a real inner product space.
- **Unitary** (operator): An isometry on a complex inner product space.
- Characterizing isometries.

Theorem 7.19. *Suppose $S \in \mathcal{L}(V)$. Then the following are equivalent.*

(a) *S is an isometry.*
(b) *$\langle Su, Sv \rangle = \langle u, v \rangle$ for all $u, v \in V$.*
(c) *$Se_1, \dots, Se_n$ is orthonormal for every orthonormal list of vectors $e_1, \dots, e_n$ in V.*
(d) *There exists an orthonormal basis $e_1, \dots, e_n$ of V such that $Se_1, \dots, Se_n$ is orthonormal.*
(e) *$S^*S = I$.*
(f) *$SS^* = I$.*
(g) *S^* is an isometry.*
(h) *S is invertible and $S^{-1} = S^*$.*

Proof. We will prove that (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (e) $\Rightarrow$ (f) $\Rightarrow$ (g) $\Rightarrow$ (h) $\Rightarrow$ (a). Let's begin.

First, suppose that S is an isometry. Let $u, v \in V$ be arbitrary. We divide into two cases (V is a real inner product space and V is a complex inner product space). If V is a real inner product space, then

$$\begin{aligned}
\langle Su, Sv \rangle &= \frac{\|Su + Sv\|^2 - \|Su - Sv\|^2}{4} && \text{Exercise 6.A.19} \\
&= \frac{\|S(u+v)\|^2 - \|S(u-v)\|^2}{4} \\
&= \frac{\|u+v\|^2 - \|u-v\|^2}{4} \\
&= \langle u, v \rangle && \text{Exercise 6.A.19}
\end{aligned}$$

as desired. On the other hand, if V is a complex vector space, then the proof is symmetric to the above except with the use of Exercise 6.A.20 instead of Exercise 6.A.19.

Second, suppose that $\langle Su, Sv \rangle = \langle u, v \rangle$ for all $u, v \in V$. Let $e_1, \dots, e_n$ be an orthonormal list of vectors in V. Then by hypothesis,

$$\langle Se_i, Se_j \rangle = \langle e_i, e_j \rangle = \delta_{ij}$$

for all $1 \leq i, j \leq n$, proving that $Se_1, \dots, Se_n$ is orthonormal, as desired.

Third, suppose that $Se_1, \dots, Se_n$ is orthonormal for every orthonormal list of vectors $e_1, \dots, e_n$ in V. By Theorem 6.14, there exists an orthonormal basis $e_1, \dots, e_n$ of V. It follows by hypothesis that $Se_1, \dots, Se_n$ is orthonormal, as desired.

Fourth, suppose that there exists an orthonormal basis $e_1, \dots, e_n$ of V such that $Se_1, \dots, Se_n$ is orthonormal. Then

$$\langle S^* Se_j, e_k \rangle = \langle Se_j, Se_k \rangle = \delta_{jk} = \langle e_j, e_k \rangle$$

for all $1 \leq j, k \leq n$. It follows that if $u, v \in V$, then

$$\begin{aligned}
\langle S^* Su, v \rangle &= \langle S^* S(a_1 e_1 + \cdots + a_n e_n), b_1 e_1 + \cdots + b_n e_n \rangle \\
&= \langle S^* S a_1 e_1, b_1 e_1 \rangle + \cdots + \langle S^* S a_n e_n, b_n e_n \rangle \\
&= \langle a_1 e_1, b_1 e_1 \rangle + \cdots + \langle a_n e_n, b_n e_n \rangle \\
&= \langle a_1 e_1 + \cdots + a_n e_n, b_1 e_1 + \cdots + b_n e_n \rangle \\
&= \langle u, v \rangle
\end{aligned}$$

Therefore, $S^* S = I$, as desired.

Fifth, suppose that $S^* S = I$. Then by Exercise 3.D.10, $SS^* = I$, as desired.

Sixth, suppose that $SS^* = I$. To prove that S^* is an isometry, it will suffice to show that $\|S^* v\| = \|v\|$ for all $v \in V$. Let $v \in V$ be arbitrary. Then

$$\|S^* v\|^2 = \langle S^* v, S^* v \rangle = \langle SS^* v, v \rangle = \langle v, v \rangle = \|v\|^2$$

Taking square roots yields the desired equality.

Seventh, suppose that S^* is an isometry. It follows by our previous chain of proofs that $(S^*)^* S^* = SS^* = I$ and $S^*(S^*)^* = S^* S = I$. Therefore, S is invertible with inverse $S^{-1} = S^*$, as desired.

Eighth, suppose that S is invertible and $S^{-1} = S^*$. Then if $v \in V$, we have that

$$\|Sv\|^2 = \langle Sv, Sv \rangle = \langle S^* Sv, v \rangle = \left\langle S^{-1} Sv, v \right\rangle = \langle v, v \rangle = \|v\|^2$$

Taking square roots yields the desired equality. ■

- It follows from (e) and (f) that every isometry is normal.

- Thus, characterizations of normal operators (e.g., the Complex Spectral Theorem) can be used to give descriptions of isometries.

Theorem 7.20. *Suppose V is a complex inner product space and $S \in \mathcal{L}(V)$. Then the following are equivalent.*

(a) *S is an isometry.*

(b) *There is an orthonormal basis of V consisting of eigenvectors of S whose corresponding eigenvalues all have absolute value 1.*

Proof. Suppose first that S is an isometry. Then by the Complex Spectral Theorem, there is an orthonormal basis $e_1, \dots, e_n$ of V consisting of the eigenvectors of S. Let $\lambda_1, \dots, \lambda_n$ be the corresponding eigenvalues. Then for each $j = 1, \dots, n$, we have that

$$|\lambda_j| = \|\lambda_j e_j\| = \|Se_j\| = \|e_j\| = 1$$

as desired.

Now suppose that there is an orthonormal basis $e_1, \dots, e_n$ of V consisting of eigenvectors of S whose corresponding eigenvalues $\lambda_1, \dots, \lambda_n$ all have absolute value 1. Let $v \in V$ be arbitrary. Then by Theorem 6.12, we have that

$$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n \qquad \|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2$$

It follows that

$$\begin{aligned} Sv &= \langle v, e_1 \rangle Se_1 + \cdots + \langle v, e_n \rangle Se_n \\ &= \langle v, e_1 \rangle \lambda_1 e_1 + \cdots + \langle v, e_n \rangle \lambda_n e_n \end{aligned}$$

Thus, we have that

$$\begin{aligned} \|Sv\|^2 &= \langle \langle v, e_1 \rangle \lambda_1 e_1 + \cdots + \langle v, e_n \rangle \lambda_n e_n, \langle v, e_1 \rangle \lambda_1 e_1 + \cdots + \langle v, e_n \rangle \lambda_n e_n \rangle \\ &= \langle \langle v, e_1 \rangle \lambda_1 e_1, \langle v, e_1 \rangle \lambda_1 e_1 \rangle + \cdots + \langle \langle v, e_n \rangle \lambda_n e_n, \langle v, e_n \rangle \lambda_n e_n \rangle \\ &= \langle v, e_1 \rangle \lambda_1 \cdot \overline{\langle v, e_1 \rangle \lambda_1} \cdot \langle e_1, e_1 \rangle + \cdots + \langle v, e_n \rangle \lambda_n \cdot \overline{\langle v, e_n \rangle \lambda_n} \cdot \langle e_n, e_n \rangle \\ &= \langle v, e_1 \rangle \lambda_1 \cdot \overline{\langle v, e_1 \rangle} \bar{\lambda}_1 \cdot 1 + \cdots + \langle v, e_n \rangle \lambda_n \cdot \overline{\langle v, e_n \rangle} \bar{\lambda}_n \cdot 1 \\ &= |\langle v, e_1 \rangle|^2 |\bar{\lambda}_1|^2 + \cdots + |\langle v, e_n \rangle|^2 |\bar{\lambda}_n|^2 \\ &= |\langle v, e_1 \rangle|^2 \cdot 1 + \cdots + |\langle v, e_n \rangle|^2 \cdot 1 \\ &= \|v\|^2 \end{aligned}$$

Taking square roots yields the desired equality. ■

7.D Polar Decomposition and Singular Value Decomposition

10/16:

- **Square root** (of a positive $T \in \mathcal{L}(V)$): The unique positive operator $R \in \mathcal{L}(V)$ such that $R^2 = T$. *Denoted by* $\sqrt{T}$.
  - The existence of such an operator is justified by Theorem 7.18.

10/18:

- Continuing with our analogy between $\mathbb{C}$ and $\mathcal{L}(V)$, we now prove an analogous theorem to the decomposition of any complex number z into the form $z = (z/|z|)|z| = (z/|z|)\sqrt{\bar{z}z}$, where $z/|z|$ (as an element of the unit circle) is analogous to an isometry, and $\bar{z}$ is analogous to the adjoint.

Theorem 7.21 (Polar Decomposition). *Suppose $T \in \mathcal{L}(V)$. Then there exists an isometry $S \in \mathcal{L}(V)$ such that*

$$T = S\sqrt{T^*T}$$

Lemma. *If $v \in V$, then*

$$\|Tv\| = \left\|\sqrt{T^*T}v\right\|$$

Proof. Let $v \in V$ be arbitrary. Then

$$\begin{aligned}\|Tv\|^2 &= \langle Tv, Tv\rangle \\ &= \langle T^*Tv, v\rangle \\ &= \left\langle \sqrt{T^*T}\sqrt{T^*T}v, v\right\rangle \\ &= \left\langle \sqrt{T^*T}v, \sqrt{T^*T}v\right\rangle \\ &= \left\|\sqrt{T^*T}v\right\|^2\end{aligned}$$

where the third equality holds because T^*T is positive by Theorem 7.17 and thus has a positive square root, and the fourth equality holds because $\sqrt{T^*T}$ is positive and thus is self-adjoint by definition. Taking square roots of the above gives the desired inequality. ■

Proof of Theorem 7.21. For this proof, we will first define a map $S_1 : \operatorname{range}\sqrt{T^*T} \to \operatorname{range} T$. We will then prove that it is a well-defined function and that it is a linear map. S_1 thus has the desired property; all that remains is to extend it to an isometry. To do so, we define $S_2 : (\operatorname{range}\sqrt{T^*T})^\perp \to (\operatorname{range} T)^\perp$ so that S, defined as S_1 on the appropriate domain and S_2 on its complement, is an isometry. Let's begin.

Let $S_1 : \operatorname{range}\sqrt{T^*Tv} \to \operatorname{range} T$ be defined by

$$S_1(\sqrt{T^*T}v) = Tv$$

for all $\sqrt{T^*T}v \in \operatorname{range}\sqrt{T^*T}$.

To prove that S_1 is a function, it will suffice to show that if $\sqrt{T^*T}v_1 = \sqrt{T^*T}v_2$, then $Tv_1 = Tv_2$. But if $\sqrt{T^*T}v_1 = \sqrt{T^*T}v_2$, then

$$\begin{aligned}\|Tv_1 - Tv_2\| &= \|T(v_1 - v_2)\| \\ &= \left\|\sqrt{T^*T}(v_1 - v_2)\right\| && \text{Lemma} \\ &= \left\|\sqrt{T^*T}v_1 - \sqrt{T^*T}v_2\right\| \\ &= 0\end{aligned}$$

Thus, by Theorem 6.2a, $Tv_1 - Tv_2 = 0$, so $Tv_1 = Tv_2$, as desired.

To prove that S_1 is a linear map, it will suffice to show that $S_1(\alpha\sqrt{T^*T}v) = \alpha S_1(\sqrt{T^*T}v)$ where $\alpha \in \mathbb{F}$ and $S_1(\sqrt{T^*T}v_1 + \sqrt{T^*T}v_2) = S_1(\sqrt{T^*T}v_1) + S_1(\sqrt{T^*T}v_2)$. But since $\sqrt{T^*T}$ and T are both linear maps themselves, we have that

$$\begin{aligned}S_1(\alpha\sqrt{T^*T}v) &= S_1(\sqrt{T^*T}(\alpha v)) \\ &= T(\alpha v) \\ &= \alpha Tv \\ &= \alpha S_1(\sqrt{T^*T}v)\end{aligned} \qquad \begin{aligned}S_1(\sqrt{T^*T}v_1 + \sqrt{T^*T}v_2) &= S_1(\sqrt{T^*T}(v_1 + v_2)) \\ &= T(v_1 + v_2) \\ &= Tv_1 + Tv_2 \\ &= S_1(\sqrt{T^*T}v_1) + S_1(\sqrt{T^*T}v_2)\end{aligned}$$

as desired.

To prove that S_1 is an isometry, it will suffice to show that $\|S_1(\sqrt{T^*T}v)\| = \|\sqrt{T^*T}v\|$ for all $\sqrt{T^*T}v \in \operatorname{range}\sqrt{T^*T}$. But if $\sqrt{T^*T}v \in \operatorname{range}\sqrt{T^*T}$, then

$$\begin{aligned}\left\|S_1(\sqrt{T^*T}v)\right\| &= \|Tv\| \\ &= \left\|\sqrt{T^*T}v\right\|\end{aligned}$$

where the first equality holds by the definition of S_1, and the second holds by the Lemma.

We now build up to our definition of S_2. For starters, notice that it follows from the Lemma that S_1 is injective much the same way it followed that S_1 was a function. Consequently, Theorem 3.4 asserts that $\operatorname{null} S_1 = \{0\}$. Thus, since $S_1 \in \mathcal{L}(\operatorname{range}\sqrt{T^*T}, \operatorname{range} T)$, we have by the Fundamental Theorem of Linear Maps that

$$\begin{aligned}\dim\operatorname{range}\sqrt{T^*T} &= \dim\operatorname{null} S_1 + \dim\operatorname{range} S_1\\ &= 0 + \dim\operatorname{range} T\\ &= \dim\operatorname{range} T\end{aligned}$$

Thus, since $\operatorname{range}\sqrt{T^*T} \subset V$ and $\operatorname{range} T \subset V$, we have that

$$\begin{aligned}\dim(\operatorname{range} T)^\perp &= \dim V - \dim\operatorname{range} T && \text{Theorem 6.21}\\ &= \dim V - \dim\operatorname{range}\sqrt{T^*T}\\ &= \dim(\operatorname{range}\sqrt{T^*T})^\perp && \text{Theorem 6.21}\end{aligned}$$

It follows that we can choose orthonormal bases $e_1, \dots, e_m$ and $f_1, \dots, f_m$ of $(\operatorname{range}\sqrt{T^*T})^\perp$ and $(\operatorname{range} T)^\perp$ of equal length. Let $S_2 : (\operatorname{range}\sqrt{T^*T})^\perp \to (\operatorname{range} T)^\perp$ be the unique linear transformation such that $Te_j = f_j$ for each $j = 1, \dots, n$ implied to exist by Theorem 3.1. Note that S_2 is also an isometry since if $x \in (\operatorname{range}\sqrt{T^*T})^\perp$, then

$$\begin{aligned}\|S_2x\|^2 &= \|S_1(a_1e_1 + \cdots + a_me_m)\|^2\\ &= \|a_1f_1 + \cdots + a_mf_m\|^2\\ &= |a_1|^2 + \cdots + |a_m|^2 && \text{Theorem 6.9}\\ &= \|a_1e_1 + \cdots + a_mf_m\|^2 && \text{Theorem 6.9}\\ &= \|x\|^2\end{aligned}$$

where taking square roots yields the desired equality.

We are now ready to define $S \in \mathcal{L}(V)$. Let $v \in V$ be arbitrary. It follows by Theorem 6.20 that we can uniquely decompose v into a sum $v = u + w$ where $u \in \operatorname{range}\sqrt{T^*T}$ and $v \in (\operatorname{range}\sqrt{T^*T})^\perp$. Thus, we define

$$Sv = S_1u + S_2w$$

We could (but will not) explicitly show based on the previously proven properties that S is well-defined and linear. We will, however, show that S is an isometry: for any $v \in V$,

$$\begin{aligned}\|Sv\|^2 &= \|S_1u + S_2v\|^2\\ &= \|S_1u\|^2 + \|S_2w\| && \text{Pythagorean Theorem}\\ &= \|u\|^2 + \|w\|^2\\ &= \|v\|^2 && \text{Pythagorean Theorem}\end{aligned}$$

Lastly, we have by its definition that for any $v \in V$,

$$(S\sqrt{T^*T})v = S(\sqrt{T^*T}v) = S_1(\sqrt{T^*T}v) = Tv$$

so $T = S\sqrt{T^*T}$, as desired. ■

10/21:

- The main conclusion from the Polar Decomposition is that *any* linear operator, no matter how ill-defined, can be decomposed into the product of an isometry and a positive operator, two very well characterized classes of operators.

  - In particular, if $\mathbb{F} = \mathbb{C}$, then T is the product of two operators, both of which are orthonormally diagonalizable (though not necessarily with respect to the same orthonormal bases).

- **Singular values** (of $T \in \mathcal{L}(V)$): The eigenvalues of $\sqrt{T^*T}$, with each value λ repeated $\dim E(\lambda, \sqrt{T^*T})$ times.
- The singular values of T are all nonnegative (because $\sqrt{T^*T}$ is a positive operator [see Theorem 7.17]).
- Each $T \in \mathcal{L}(V)$ has $\dim V$ singular values (because $\sqrt{T^*T}$ is positive, hence self-adjoint, hence $\sqrt{T^*T}$ has a diagonal matrix [see the Real Spectral Theorem], hence $\sqrt{T^*T}$ has $\dim V$ distinct eigenvalues).
- We now show that every operator on V can be described in terms of its singular values and two orthonormal bases on V.

Theorem 7.22 (Singular Value Decomposition)**.** *Suppose $T \in \mathcal{L}(V)$ has singular values $s_1, \dots, s_n$. Then there exist orthonormal bases $e_1, \dots, e_n$ and $f_1, \dots, f_n$ of V such that*

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n$$

for every $v \in V$.

Proof. Applying the Real Spectral Theorem to the self-adjoint operator $\sqrt{T^*T}$ reveals that V has an orthonormal basis $e_1, \dots, e_n$ of eigenvectors of $\sqrt{T^*T}$. Therefore, if we let $v \in V$ be arbitrary, then we have that

$$\begin{aligned}
Tv &= (S\sqrt{T^*T})v && \text{Polar Decomposition} \\
&= S(\sqrt{T^*T}(\langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n)) && \text{Theorem 6.12} \\
&= S(\langle v, e_1 \rangle \sqrt{T^*T}e_1 + \cdots + \langle v, e_n \rangle \sqrt{T^*T}e_n) \\
&= S(\langle v, e_1 \rangle s_1 e_1 + \cdots + \langle v, e_n \rangle s_n e_n) \\
&= s_1 \langle v, e_1 \rangle Se_1 + \cdots + s_n \langle v, e_n \rangle Se_n \\
&= s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n && \text{Theorem 7.19}
\end{aligned}$$

where $s_1, \dots, s_n$ are the singular values of T (the eigenvalues of $\sqrt{T^*T}$) and $f_1, \dots, f_n$ is another orthonormal basis of V. ∎

- If $e_1, \dots, e_n$ and $f_1, \dots, f_n$ are orthonormal bases of V that satisfy the Singular Value Decomposition for some operator T, then $Te_j = s_j f_j$ for each $j = 1, \dots, n$.
  - In other words, every operator on V has a diagonal matrix with respect to some orthonormal *bases* (plural) of V.
- The Singular Value Decomposition has many applications, especially in the realm of computational linear algebra, where working with T^*T is much easier than working with $\sqrt{T^*T}$. A powerful tool in this pursuit is the following.

Theorem 7.23. *Suppose $T \in \mathcal{L}(V)$. Then the singular values of T are the nonnegative square roots of the eigenvalues of T^*T, with each eigenvalue λ repeated* $\dim E(\lambda, T^*T)$ *times.*

Proof. Since T^*T is positive and self-adjoint, we have by the hyperref[trm:RealSpectral]Real Spectral Theorem that there exists an orthonormal basis $e_1, \dots, e_n$ of V and nonnegative numbers $\lambda_1, \dots, \lambda_n$ such that $T^*Te_j = \lambda_j e_j$ for each $j = 1, \dots, n$. It follows since $\sqrt{T^*T}$ is also a positive, self-adjoint operator that its eigenvalues (which exist) must be of the form $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$ to satisfy $(\sqrt{T^*T})^2 = T^*T$ and to be nonnegative. ∎