# Section 7

# The adjoint of a transformation and self-adjoint transformations

Throughout this section, $V$ is a finite-dimensional inner product space over a field $F$ (where, as before, $F = \mathbb{R}$ or $\mathbb{C}$) with inner product $\langle \cdot, \cdot \rangle$.

**Definition 7.1** Let $T \colon V \to V$ be a linear transformation. The *adjoint* of $T$ is a map $T^* \colon V \to V$ such that

$$\langle T(v), w \rangle = \langle v, T^*(w) \rangle \qquad \text{for all } v, w \in V.$$

**Remark:** More generally, if $T \colon V \to W$ is a linear map between inner product spaces, the adjoint $T^* \colon W \to V$ is a map satisfying the above equation for all $v \in V$ and $w \in W$. Appropriate parts of what we describe here can be done in this more general setting.

**Lemma 7.2** *Let $V$ be a finite-dimensional inner product space and let $T \colon V \to V$ be a linear transformation. Then there is a unique adjoint $T^*$ for $T$ and, moreover, $T^*$ is a linear transformation.*

Proof: We first show that if $T^*$ exists, then it is unique. For if $S \colon V \to V$ also satisfies the same condition, then

$$\langle v, T^*(w) \rangle = \langle T(v), w \rangle = \langle v, S(w) \rangle$$

for all $v, w \in V$. Hence

$$\langle v, T^*(w) \rangle - \langle v, S(w) \rangle = 0,$$

that is,

$$\langle v, T^*(w) - S(w) \rangle = 0 \qquad \text{for all } v, w \in V.$$

Let us fix $w \in V$ and take $v = T^*(w) - S(w)$. Then

$$\langle T^*(w) - S(w), T^*(w) - S(w) \rangle = 0.$$

The axioms of an inner product space tell us

$$T^*(w) - S(w) = \mathbf{0}$$

so

$$S(w) = T^*(w) \qquad \text{for all } w \in V,$$

as claimed.

It remains to show that such a linear map $T^*$ actually exists. Let $\mathscr{B} = \{e_1, e_2, \dots, e_n\}$ be an orthonormal basis for $V$. (The Gram–Schmidt Process guarantees that this exists.) Let $A = [\alpha_{ij}]$ be the matrix of $T$ with respect to $\mathscr{B}$. Define $T^* \colon V \to V$ be the linear map whose matrix is the conjugate transpose of $A$ with respect to $\mathscr{B}$. Thus

$$T^*(e_j) = \sum_{i=1}^{n} \bar{\alpha}_{ji} e_i \qquad \text{for } j = 1,\, 2,\, \dots,\, n.$$

(Here we are using Proposition 2.7 to guarantee that this determines a unique linear transformation $T^*$.) Note also that

$$T(e_j) = \sum_{i=1}^{n} \alpha_{ij} e_i \qquad \text{for } j = 1,\, 2,\, \dots,\, n.$$

**Claim:** $\langle T(v), w \rangle = \langle v, T^*(w) \rangle$ for all $v, w \in V$.

Write $v = \sum_{j=1}^{n} \beta_j e_j$ and $w = \sum_{k=1}^{n} \gamma_k e_k$ in terms of the basis $\mathscr{B}$. Then

$$\begin{aligned}
\langle T(v), w \rangle &= \left\langle T\left( \sum_{j=1}^{n} \beta_j e_j \right), \sum_{k=1}^{n} \gamma_k e_k \right\rangle \\
&= \left\langle \sum_{j=1}^{n} \beta_j T(e_j), \sum_{k=1}^{n} \gamma_k e_k \right\rangle \\
&= \left\langle \sum_{j=1}^{n} \beta_j \sum_{i=1}^{n} \alpha_{ij} e_i, \sum_{k=1}^{n} \gamma_k e_k \right\rangle \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} \sum_{k=1}^{n} \beta_j \alpha_{ij} \bar{\gamma}_k \langle e_i, e_k \rangle \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} \beta_j \alpha_{ij} \bar{\gamma}_i,
\end{aligned}$$

while

$$\begin{aligned}
\langle v, T^*(w)\rangle &= \left\langle \sum_{j=1}^{n} \beta_j e_j, T^*\left(\sum_{k=1}^{n} \gamma_k e_k\right)\right\rangle \\
&= \left\langle \sum_{j=1}^{n} \beta_j e_j, \sum_{k=1}^{n} \gamma_k T^*(e_k)\right\rangle \\
&= \left\langle \sum_{j=1}^{n} \beta_j e_j, \sum_{k=1}^{n} \gamma_k \sum_{i=1}^{n} \bar{a}_{ki} e_i\right\rangle \\
&= \sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{i=1}^{n} \beta_j \bar{\gamma}_k a_{ki} \langle e_j, e_i\rangle \\
&= \sum_{j=1}^{n}\sum_{k=1}^{n} \beta_j \bar{\gamma}_k a_{kj} \\
&= \langle T(v), w\rangle.
\end{aligned}$$

Hence $T^*$ is indeed the adjoint of $T$. □

We also record what was observed in the course of this proof:

> If $A = \mathrm{Mat}_{\mathscr{B},\mathscr{B}}(T)$ is the matrix of $T$ with respect to an *orthonormal basis*, then
>
> $$\mathrm{Mat}_{\mathscr{B},\mathscr{B}}(T^*) = \bar{A}^{\mathrm{T}}$$
>
> (the *conjugate transpose* of $A$).

**Definition 7.3** A linear transformation $T\colon V \to V$ is *self-adjoint* if $T^* = T$.

Interpreting this in terms of the matrices (using our observation above), we conclude:

**Lemma 7.4** (i) *A real matrix $A$ defines a self-adjoint transformation if and only if it is* symmetric*: $A^{\mathrm{T}} = A$.*

(ii) *A complex matrix $A$ defines a self-adjoint transformation if and only if it is* Hermitian*: $\bar{A}^{\mathrm{T}} = A$.* □

The most important theorem concerning self-adjoint transformation is the following:

**Theorem 7.5** *A self-adjoint transformation of a finite-dimensional inner product space is diagonalisable.*

Interpreting this in terms of matrices gives us:

**Corollary 7.6** (i) *A real symmetric matrix is diagonalisable.*

(ii) *A Hermitian matrix is diagonalisable.*

We finish the course by establishing Theorem 7.5. First we establish the main tools needed to prove that result.

**Lemma 7.7** *Let $V$ be a finite-dimensional inner product space and $T \colon V \to V$ be a self-adjoint transformation. Then the characteristic polynomial is a product of linear factors and every eigenvalue of $T$ is real.*

PROOF: Any polynomial is factorisable over $\mathbb{C}$ into a product of linear factors. Thus it is sufficient to show all the roots of the characteristic polynomial are real.

Let $W$ be an inner product space over $\mathbb{C}$ with the same dimension as $V$ and let $S \colon W \to W$ be a linear transformation whose matrix $A$ with respect to an orthonormal basis is the same as that of $T$ with respect to an orthonormal basis for $V$. Then $S$ is also self-adjoint since $\bar{A}^{\mathrm{T}} = A$ (because $T^* = T$). (Essentially this process deals with the fact that $V$ might be a vector space over $\mathbb{R}$, so we replace it by one over $\mathbb{C}$ that in all other ways is the same.)

Let $\lambda \in \mathbb{C}$ be a root of $c_S(x) = \det(xI - A) = c_T(x)$. Then $\lambda$ is an eigenvalue of $S$, so there exists an eigenvector $v \in W$ for $S$:

$$S(v) = \lambda v.$$

Therefore

$$\langle S(v), v \rangle = \langle \lambda v, v \rangle = \lambda \|v\|^2,$$

but also

$$\langle S(v), v \rangle = \langle v, S^*(v) \rangle = \langle v, S(v) \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \|v\|^2.$$

Hence

$$\lambda \|v\|^2 = \bar{\lambda} \|v\|^2$$

and since $v \neq \mathbf{0}$, we conclude $\lambda = \bar{\lambda}$. This shows that $\lambda \in \mathbb{R}$ and the lemma is proved. □

**Lemma 7.8** *Let $V$ be an inner product space and $T \colon V \to V$ be a linear map. If $U$ is a subspace of $V$ such that $T(U) \subseteq U$ (i.e., $U$ is $T$-invariant), then $T^*(U^\perp) \subseteq U^\perp$ (i.e., $U^\perp$ is $T^*$-invariant).*

PROOF: Let $v \in U^\perp$. Then for any $u \in U$, we have

$$\langle u, T^*(v) \rangle = \langle T(u), v \rangle = 0,$$

since $T(u) \in U$ (by assumption) and $v \in U^\perp$. Hence $T^*(v) \in U^\perp$. □

These two lemmas now enable us to prove the main theorem about diagonalisation of self-adjoint transformations.

PROOF OF THEOREM 7.5: We proceed by induction on $n = \dim V$. If $n = 1$, then $T$ is represented by a $1 \times 1$ matrix, which is already diagonal.

Consider the characteristic polynomial $c_T(x)$. By Lemma 7.7, this is a product of linear factors. In particular, there exists some root $\lambda \in F$. Let $v_1$ be an eigenvector with eigenvalue $\lambda$. Let $U = \mathrm{Span}(v_1)$ be the 1-dimensional subspace spanned by $v_1$. By Theorem 6.16,

$$V = U \oplus U^{\perp}.$$

Now as $T(v_1) = \lambda v_1 \in U$, we see that $U$ is $T$-invariant. Hence $U^{\perp}$ is also $T$-invariant by Lemma 7.8 (since $T^* = T$).

Now consider the restriction $S = T|_{U^{\perp}} : U^{\perp} \to U^{\perp}$ of $T$ to $U^{\perp}$. This is self-adjoint, since

$$\langle T(v), w \rangle = \langle v, T(w) \rangle \qquad \text{for all } v, w, \in U^{\perp}$$

tells us

$$(T|_{U^{\perp}})^* = T|_{U^{\perp}}.$$

By induction, $S = T|_{U^{\perp}}$ is diagonalisable. Hence there is a basis $\{v_2, \dots, v_n\}$ for $U^{\perp}$ of eigenvectors for $T$. Then as $V = U \oplus U^{\perp}$, we conclude that $\{v_1, v_2, \dots, v_n\}$ is a basis for $V$ consisting of eigenvectors for $T$. Hence $T$ is diagonalisable and the proof is complete. □