□

14.4 Linear Functionals and Adjoint Operators

When we consider a linear operator T on a vector space V equipped with an inner product, we can construct a related but generally different linear operator $T^\dagger$ on V called the **adjoint** of T. When the adjoint $T^\dagger$ happens to be equal to T, the operator is said to be Hermitian. To understand adjoints, we first need to develop the concept of a linear functional.

A **linear functional** ϕ on the vector space V is a linear map from V to the numbers $\mathbb{F}$: for $v \in V$, $\phi(v) \in \mathbb{F}$ (as usual, $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$). A linear functional has the following two properties:

1. $\phi(v_1 + v_2) = \phi(v_1) + \phi(v_2)$, with $v_1, v_2 \in V$.
2. $\phi(av) = a\phi(v)$ for $v \in V$ and $a \in \mathbb{F}$.

Example 14.8. *Linear functional on* $\mathbb{R}^3$.

Consider the real vector space $\mathbb{R}^3$ with the inner product equal to the familiar dot product. For any $v = (v_1, v_2, v_3) \in \mathbb{R}^3$, we define the functional ϕ as follows:

$$\phi(v) = 3v_1 + 2v_2 - 4v_3. \tag{14.4.1}$$

Linearity follows because the components v_1, v_2, and v_3 appear linearly on the right-hand side. We can actually use the vector $u = (3, 2, -4)$ to write the linear functional as an inner product. Indeed, one can readily see that for any $v \in \mathbb{R}^3$,

$$\phi(v) = \langle u, v\rangle = \langle (3,2,-4), (v_1, v_2, v_3)\rangle = 3v_1 + 2v_2 - 4v_3. \tag{14.4.2}$$

This is no accident; we will now prove that any linear functional $\phi(v)$ on a vector space V admits such representation with some suitable choice of vector u.

□

Theorem 14.4.1. *Let ϕ be a linear functional on V. There is a unique vector $u \in V$ such that $\phi(v) = \langle u, v\rangle$ for all $v \in V$.*

Proof. Consider an orthonormal basis $(e_1, \ldots, e_n)$, and write the vector v as

$$v = \langle e_1, v\rangle e_1 + \cdots + \langle e_n, v\rangle e_n. \tag{14.4.3}$$

When ϕ acts on v, we find

$$\begin{aligned} \phi(v) &= \phi\left(\langle e_1, v\rangle e_1 + \cdots + \langle e_n, v\rangle e_n\right) \\ &= \langle e_1, v\rangle \phi(e_1) + \cdots + \langle e_n, v\rangle \phi(e_n) \\ &= \langle \phi(e_1)^* e_1, v\rangle + \cdots + \langle \phi(e_n)^* e_n, v\rangle \\ &= \left\langle \phi(e_1)^* e_1 + \cdots + \phi(e_n)^* e_n, v\right\rangle, \end{aligned} \tag{14.4.4}$$

where we first used linearity and then conjugate homogeneity to bring the constants $\phi(e_i)$ inside the inner products. We have thus shown that, as claimed,

$$\phi(v) = \langle u, v\rangle \quad \text{with} \quad u = \phi(e_1)^* e_1 + \cdots + \phi(e_n)^* e_n. \tag{14.4.5}$$

Next, we prove that this u is unique. If there exists another vector u' that also gives the correct result for all v, then $\langle u', v\rangle = \langle u, v\rangle$, which implies $\langle u-u', v\rangle = 0$ for all v. Taking $v = u' - u$, we see that this implies $u' - u = 0$, or $u' = u$, proving uniqueness. This proof applies for finite-dimensional vector spaces. The result, however, is true for infinite-dimensional vector spaces when ϕ is what is called a *continuous* linear functional.

□

We can now address the construction of the adjoint. Consider a linear operator T and a functional $\phi(v)$ defined as follows:

$$\phi(v) = \langle u,\ Tv \rangle. \tag{14.4.6}$$

This is clearly a linear functional, whatever the operator T is. Since any linear functional can be written as $\langle w,\ v \rangle$, with some suitable vector w, we write

$$\langle u,\ Tv \rangle \ = \ \langle \#, v \rangle. \tag{14.4.7}$$

Of course, the vector # must depend on the vector u that appears on the left-hand side. Moreover, it must have something to do with the operator T, which does not appear on the right-hand side. We can think of # as a function of the vector u and thus write $\# = T^\dagger u$, where $T^\dagger$ denotes a function, not obviously linear, from V to V. So we think of $T^\dagger u$ as the vector obtained by acting with some function $T^\dagger$ on u. The above equation is written as

$$\langle u, Tv \rangle \ = \ \langle T^\dagger u, v \rangle. \tag{14.4.8}$$

Our next step is to show that, in fact, $T^\dagger$ is a linear operator on V.

Theorem 14.4.2. $T^\dagger \in \mathcal{L}(V)$.

Proof. For this purpose, consider

$$\langle u_1 + u_2, Tv \rangle \ = \ \langle T^\dagger (u_1 + u_2), v \rangle \tag{14.4.9}$$

with u_1, u_2, and v arbitrary vectors. Expand the left-hand side to get

$$\begin{aligned} \langle u_1 + u_2, Tv \rangle \ &= \ \langle u_1, Tv \rangle + \langle u_2, Tv \rangle \\ &= \ \langle T^\dagger u_1, v \rangle + \langle T^\dagger u_2, v \rangle \\ &= \ \left\langle T^\dagger u_1 + T^\dagger u_2, v \right\rangle. \end{aligned} \tag{14.4.10}$$

Comparing the right-hand sides of the last two equations, we get the desired

$$T^\dagger(u_1+u_2) \;=\; T^\dagger u_1 + T^\dagger u_2. \tag{14.4.11}$$

Having established linearity, we now establish homogeneity. Consider

$$\langle au,\, Tv\rangle \;=\; \langle T^\dagger(au),\, v\rangle, \tag{14.4.12}$$

with $a \in \mathbb{C}$, and u and v arbitrary vectors. The left-hand side is

$$\langle au, Tv\rangle \;=\; a^*\langle u, Tv\rangle \;=\; a^*\langle T^\dagger u, v\rangle \;=\; \langle aT^\dagger u, v\rangle. \tag{14.4.13}$$

This time we conclude that

$$T^\dagger(au) \;=\; a\,T^\dagger u. \tag{14.4.14}$$

This completes the proof that $T^\dagger$, so defined, is a linear operator on V.

□

The operator $T^\dagger \in \mathcal{L}(V)$ is called the **adjoint** of T. Its operational definition is the relation

$$\boxed{\langle u, Tv\rangle \;=\; \langle T^\dagger u, v\rangle.} \tag{14.4.15}$$

A couple of important properties are readily proven. The first is

$$(ST)^\dagger = T^\dagger S^\dagger. \tag{14.4.16}$$

To see this, apply the operational definition twice: $\langle u,\, STv\rangle = \langle S^\dagger u,\, Tv\rangle = \langle T^\dagger S^\dagger u,\, v\rangle$. By definition, $\langle u,\, STv\rangle = \langle (ST)^\dagger u,\, v\rangle$. Comparison leads to the claimed relation. The second property states that the adjoint of the adjoint of an operator is the original operator:

$$(S^\dagger)^\dagger = S. \tag{14.4.17}$$

To see this, first consider $\langle u,\, S^\dagger v\rangle = \langle (S^\dagger)^\dagger u,\, v\rangle$. But $\langle u,\, S^\dagger v\rangle = \langle S^\dagger v,\, u\rangle^* = \langle v,\, Su\rangle^* = \langle Su,\, v\rangle$. Comparing with the first result, we have shown that $(S^\dagger)^\dagger u = Su$, for any u. This is the content of the claimed relation. It is a direct result of $(S^\dagger)^\dagger = S$ that $\langle Su,\, v\rangle = \langle u,\, S^\dagger v\rangle$, as you can see by reading this equation from right to left. In summary, an operator can be moved from the left input of the inner product to the right input, and vice versa, by adding a dagger.

Example 14.9. *Adjoint operator on* $\mathbb{C}^3$.
Let $v = (v_1, v_2, v_3)$, with $v_i \in \mathbb{C}$, denote a vector in the three-dimensional complex vector space $\mathbb{C}^3$. Define a linear operator T that acts on v as follows:

$$T(v_1, v_2, v_3) = (2v_2 + iv_3,\ v_1 - iv_2,\ 3iv_1 + v_2 + 7v_3). \tag{14.4.18}$$

Assume the inner product is the standard one on $\mathbb{C}^3$: $\langle u, v\rangle = u_1^* v_1 + u_2^* v_2 + u_3^* v_3$. We want to find the action of $T^\dagger$ on a vector and wish to determine the matrix representations of T and $T^\dagger$ using the orthonormal basis $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, $e_3 = (0, 0, 1)$.

We introduce the vector $u = (u_1, u_2, u_3)$ and use the basic identity $\langle Tu, v\rangle = \langle u, T^\dagger v\rangle$. The left-hand side of the identity gives

$$\langle Tu, v\rangle = (2u_2 + iu_3)^* v_1 + (u_1 - iu_2)^* v_2 + (3iu_1 + u_2 + 7u_3)^* v_3. \tag{14.4.19}$$

To identify this with $\langle u, T^\dagger v\rangle$, we rewrite the right-hand side, factoring the various u_i^*'s:

$$\langle u, T^\dagger v\rangle = u_1^*(v_2 - 3iv_3) + u_2^*(2v_1 + iv_2 + v_3) + u_3^*(-iv_1 + 7v_3). \tag{14.4.20}$$

We can therefore read the desired action of $T^\dagger$:

$$T^\dagger(v_1, v_2, v_3) = (v_2 - 3iv_3,\ 2v_1 + iv_2 + v_3,\ -iv_1 + 7v_3). \tag{14.4.21}$$

To find the matrix representations, think of (14.4.18) as follows:

$$\begin{pmatrix} * & * & * \\ * & * & * \\ * & * & * \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 2v_2 + iv_3 \\ v_1 - iv_2 \\ 3iv_1 + v_2 + 7v_3 \end{pmatrix}. \tag{14.4.22}$$

The matrix on the left-hand side is the representation of T. We immediately see that

$$T = \begin{pmatrix} 0 & 2 & i \\ 1 & -i & 0 \\ 3i & 1 & 7 \end{pmatrix}, \qquad T^\dagger = \begin{pmatrix} 0 & 1 & -3i \\ 2 & i & 1 \\ -i & 0 & 7 \end{pmatrix}, \tag{14.4.23}$$

using (14.4.21) to write the matrix for $T^\dagger$. These matrices are related: one is the transpose and complex conjugate of the other. This is not an accident, as we will see now.

□

Let us calculate adjoints using matrix notation. Let $u = e_i$ and $v = e_j$ where e_i and e_j are *orthonormal* basis vectors. Then the definition $\langle T^\dagger u, v\rangle = \langle u, Tv\rangle$ can be written (with repeated indices summed) as

$$\begin{aligned} \langle T^\dagger e_i, e_j\rangle &= \langle e_i, Te_j\rangle, \\ \langle (T^\dagger)_{ki} e_k, e_j\rangle &= \langle e_i, T_{kj} e_k\rangle, \\ ((T^\dagger)_{ki})^* \, \delta_{kj} &= T_{kj}\, \delta_{ik}, \\ ((T^\dagger)_{ji})^* &= T_{ij}. \end{aligned} \tag{14.4.24}$$

Relabeling i and j and taking the complex conjugate, we find the familiar relation between a matrix and its adjoint:

$$\boxed{\text{In an orthonormal basis,} \quad (T^\dagger)_{ij} = (T_{ji})^*.} \tag{14.4.25}$$

The adjoint matrix is the transpose and complex conjugate matrix *as long* as we use an orthonormal basis. If we did not, in the equation above, $\langle e_i, e_j\rangle = \delta_{ij}$ would be replaced by $\langle e_i, e_j\rangle = g_{ij}$, where g_{ij} is some constant matrix that would appear in the rule for the construction of the adjoint matrix.

Exercise 14.3. *Let $(e_1, \cdots, e_n)$ be a basis with inner product $\langle e_i, e_j\rangle = m_i\delta_{ij}$(i not summed), where $m_i > 0$ for all i. Show that*

$$(T^\dagger)_{ij} = \frac{m_j}{m_i}\, T^*_{ji}. \tag{14.4.26}$$