14.5 Hermitian and Unitary Operators

Before we begin looking at special kinds of operators, let us consider a very surprising fact about operators on complex vector spaces. Suppose we have an operator T such that for any vector $v \in V$ the following inner product vanishes:

$$\langle v, Tv \rangle = 0 \quad \text{for all } v \in V. \tag{14.5.1}$$

What can we say about the operator T? The condition states that T is an operator that acting on any vector v gives a vector orthogonal to v. In a two-dimensional *real* vector space, this could be the operator that rotates any vector by ninety degrees, a nontrivial operator. It is quite surprising and important that for *complex* vector spaces any such operator *necessarily* vanishes. This is a theorem:

Theorem 14.5.1. *Let T be a linear operator in a complex vector space V:*

$$\boxed{\text{If } \langle v, Tv\rangle = 0 \text{ for all } v \in V, \text{ then } T = 0.} \tag{14.5.2}$$

Proof. Any proof must be such that it fails to work for a real vector space. Note that the result follows if we can prove that $\langle u, Tv\rangle = 0$, for all $u, v \in V$. Indeed, if this holds then take $u = Tv$. Then $\langle Tv, Tv\rangle = 0$ for all v implies that $Tv = 0$ for all v, and therefore $T = 0$.

We will thus try to show that $\langle u, Tv\rangle = 0$ for all $u, v \in V$. All we know is that objects of the form $\langle \#, T\#\rangle$ vanish, whatever # is. So we must aim to form linear combinations of such terms in order to reproduce $\langle u, Tv\rangle$. We begin by trying the following:

$$\langle u+v, T(u+v)\rangle \;-\; \langle u-v, T(u-v)\rangle \;=\; 2\langle u, Tv\rangle \;+\; 2\langle v, Tu\rangle. \tag{14.5.3}$$

We see that the "diagonal" terms vanish, but instead of getting just $\langle u, Tv\rangle$, we also get $\langle v, Tu\rangle$. Here is where complex numbers help. We can get the same two terms but with opposite signs as follows:

$$\langle u+iv, T(u+iv)\rangle \;-\; \langle u-iv, T(u-iv)\rangle \;=\; 2i\langle u, Tv\rangle \;-\; 2i\,\langle v, Tu\rangle. \tag{14.5.4}$$

In checking this don't forget that, for example, $\langle iv, u\rangle = -i\langle v, u\rangle$. It follows from the last two relations that

$$\begin{aligned}\langle u, Tv\rangle \;=\; \tfrac{1}{4}\Big(&\langle u+v, T(u+v)\rangle - \langle u-v, T(u-v)\rangle \\ &+ \tfrac{1}{i}\langle u+iv, T(u+iv)\rangle - \tfrac{1}{i}\langle u-iv, T(u-iv)\rangle\Big).\end{aligned} \tag{14.5.5}$$

The condition $\langle v, Tv\rangle = 0$ for all v implies that each term on the above right-hand side vanishes, thus showing that $\langle u, Tv\rangle = 0$ for all $u, v \in V$. As explained above this proves the result.

□

Exercise 14.4. *If a nonvanishing operator T existed that acting on any vector v on a complex vector space gave a vector orthogonal to v, it would contradict the above theorem. Give an independent reason why such an operator cannot exist (Hint: if nothing comes to mind, reread example 13.12).*

An operator T is said to be **Hermitian** if $T^\dagger = T$. Hermitian operators are pervasive in quantum mechanics. The above theorem in fact helps us discover Hermitian operators. You recall that the expectation value of a Hermitian operator, on any state, is real. It is also true, however, that any operator whose expectation value is real for all states must be Hermitian:

$$\boxed{T = T^\dagger \text{ if and only if } \langle v, Tv\rangle \in \mathbb{R} \text{ for all } v.} \tag{14.5.6}$$

To prove this we first go from left to right. If $T = T^\dagger$, then

$$\langle v, Tv\rangle = \langle T^\dagger v, v\rangle = \langle Tv, v\rangle = \langle v, Tv\rangle^*, \tag{14.5.7}$$

showing that $\langle v, Tv\rangle$ is real. To go from right to left, first note that the reality condition means that

$$\langle v, Tv\rangle = \langle Tv, v\rangle = \langle v, T^\dagger v\rangle. \tag{14.5.8}$$

Now the leftmost and rightmost terms can be combined to give $\langle v, (T - T^\dagger)v\rangle = 0$, which holding for all v implies, by the theorem, that $T = T^\dagger$.

We have shown (theorem 13.6.2) that on a complex vector space any linear operator has at least one eigenvalue. Below we show that the eigenvalues of a Hermitian operator are real numbers. Moreover, while eigenvectors corresponding to different eigenvalues are in general linearly independent, for Hermitian operators they are guaranteed to be orthogonal. These results were established in section 5.3 for state spaces of wave functions with the obvious inner product arising from integration. Here the inner product is completely general, and so is the state space.

Theorem 14.5.2. *The eigenvalues of Hermitian operators are real.*

Proof. Let v be a nonzero eigenvector of the Hermitian operator T with eigenvalue λ: $Tv = \lambda v$. Taking the inner product with v, we find that

$$\langle v, Tv\rangle = \langle v, \lambda v\rangle = \lambda\langle v, v\rangle. \tag{14.5.9}$$

Since T is Hermitian, we can also evaluate $\langle v, Tv\rangle$ as follows:

$$\langle v, Tv\rangle = \langle Tv, v\rangle = \langle \lambda v, v\rangle = \lambda^* \langle v, v\rangle. \tag{14.5.10}$$

The above equations give $(\lambda - \lambda^*)\langle v, v\rangle = 0$, and since v is not the zero vector, we conclude that $\lambda^* = \lambda$, showing that λ is real.

□

Theorem 14.5.3. *Eigenvectors of a Hermitian operator associated with different eigenvalues are orthogonal.*

Proof. Let v_1 and v_2 be eigenvectors of the operator T:

$$Tv_1 = \lambda_1 v_1, \qquad Tv_2 = \lambda_2 v_2, \tag{14.5.11}$$

with λ_1 and λ_2 real, by Theorem 14.5.2, and different from each other. Consider the inner product $\langle v_2, Tv_1\rangle$, and evaluate it in two different ways by following the direction of the arrows emanating from the central term:

$$\lambda_2\langle v_2, v_1\rangle = \langle \lambda_2 v_2, v_1\rangle = \langle Tv_2, v_1\rangle \underset{\longleftarrow}{=} \langle v_2, Tv_1\rangle \underset{\longrightarrow}{=} \langle v_2, \lambda_1 v_1\rangle = \lambda_1\langle v_2, v_1\rangle. \tag{14.5.12}$$

Going left, we used the Hermiticity of T. Equating the leftmost and rightmost terms, we find that

$$(\lambda_1 - \lambda_2)\langle v_1, v_2\rangle = 0. \tag{14.5.13}$$

The assumption $\lambda_1 \neq \lambda_2$ leads to the claimed orthogonality: $\langle v_1, v_2\rangle = 0$.

□

We can finally give a rather simple characterization of orthogonal projectors.

Theorem 14.5.4. *An operator P such that $P^2 = P$ and $P^\dagger = P$ is an orthogonal projector.*

Proof. For any $v \in V$, we can write

$$v = (v - Pv) + Pv, \quad v - Pv \in \text{null}\, P, \quad \text{and} \quad Pv \in \text{range}\, P. \tag{14.5.14}$$

Any vector both in the null space of P and in the range of P must be zero. Indeed, let $x \in$ range P be nonzero. This means there is a $y \in V$, nonzero, such that $x = Py$. From this it follows that $Px = P^2y = Py = x$, demonstrating that x cannot be in the null space of P. All of this shows that $P^2 = P$ implies that

$$V = \text{range}\, P \oplus \text{null}\, P. \tag{14.5.15}$$

Now it remains to be shown that the range and null spaces above are orthogonal. Let $u \in$ range P and $w \in$ null P. Since P leaves u invariant, is Hermitian, and kills w, we have

$$\langle u, w\rangle = \langle Pu, w\rangle = \langle u, Pw\rangle = 0, \tag{14.5.16}$$

proving the desired orthonormality.

□

Let us now consider another important class of linear operators on a complex vector space, the so-called unitary operators. An operator $U \in \mathcal{L}(V)$ in a complex vector space V is said to be a **unitary operator** if it is surjective and does not change the norm of the vector it acts upon:

$$\|Uv\| = \|v\|\,, \text{ for all } v \in V. \tag{14.5.17}$$

Note that U can only kill vectors of zero length, and since the only such vector is the zero vector, null $U = 0$ and U is injective. By including the condition of surjectivity, we tailored the definition to be useful even for infinite-dimensional spaces. In finite-dimensional vector spaces, injectivity implies surjectivity and invertibility. In infinite-dimensional vector spaces, however, invertibility requires both injectivity and surjectivity. Since U is also assumed to be surjective, a unitary operator U is always invertible.

Example 14.10. *A unitary operator proportional to the identity.*
The operator $\lambda\mathbb{1}$ with λ a complex number of unit norm is unitary. Indeed, with $|\lambda| = 1$ we have $\|\lambda\mathbb{1}v\| = \|\lambda v\| = |\lambda|\|v\| = \|v\|$ for all $v \in V$, showing the operator preserves the length of all vectors. Moreover, the operator is clearly surjective since for any $v \in V$ we have $v = (\lambda\mathbb{1})\frac{1}{\lambda}v$.

□

For another useful characterization of unitary operators, we begin by squaring the invariance of the norm condition (14.5.17):

$$\langle Uu, Uu\rangle = \langle u, u\rangle. \tag{14.5.18}$$

By the definition of adjoint,

$$\langle u, U^\dagger U\, u\rangle = \langle u, u\rangle \;\Rightarrow\; \langle u,\, (U^\dagger U - \mathbb{1})u\rangle \;=\; 0 \;\text{ for all }\; u. \tag{14.5.19}$$

Theorem 14.5.1 then implies $U^\dagger U = \mathbb{1}$. Since U is invertible, $U^\dagger$ is the inverse of U, and we also have $UU^\dagger = \mathbb{1}$:

$$\boxed{U \text{ is unitary } \Longleftrightarrow\; U^\dagger U = UU^\dagger = \mathbb{1}.} \tag{14.5.20}$$

The right-to-left arrow holds because any operator U that obeys these identities is unitary (invertible and norm preserving). Unitary operators also preserve *inner products*:

$$\langle\, Uu,\; Uv\rangle \;=\; \langle u, v\rangle. \tag{14.5.21}$$

This follows immediately by moving the second U to act on the first input and using $U^\dagger U = \mathbb{1}$.

Example 14.11. *From Hermitian to unitary operators.*
We noted earlier that the exponentiation of certain Hermitian operators gives interesting operators. In section 10.1, for example, we considered the momentum operator, which is Hermitian, multiplied it by i as well as some real constants, and exponentiated it to obtain a translation operator. We did similarly with the angular momentum operator $\hat{L}_z$, a Hermitian operator that multiplied by i and a real constant is exponentiated to give a rotation operator. We said that the momentum operator generates translations, and the angular momentum operator generates rotations. The general rule was noted in exercise 10.2:

$$\boxed{\hat{M} \text{ Hermitian } \Rightarrow\; e^{i\hat{M}} \;\text{ unitary.}} \tag{14.5.22}$$

This is an important result, so let us make sure it is completely clear. Since $(\hat{A}_1 + \hat{A}_2)^\dagger = \hat{A}_1^\dagger + \hat{A}_2^\dagger$ and $\mathbb{1}^\dagger = \mathbb{1}$, the series expansion of the exponential implies that $(e^{\hat{A}})^\dagger = e^{\hat{A}^\dagger}$. It is then clear that for a Hermitian $\hat{M}$,

$$\left(e^{i\hat{M}}\right)^\dagger = e^{(i\hat{M})^\dagger} = e^{-i\hat{M}}. \tag{14.5.23}$$

It then follows that

$$\left(e^{i\hat{M}}\right)^\dagger e^{i\hat{M}} = e^{-i\hat{M}} e^{i\hat{M}} = \mathbb{1}, \tag{14.5.24}$$

confirming that $e^{i\hat{M}}$ is a unitary operator, as its Hermitian conjugate is its inverse. In quantum mechanics, unitary operators act naturally on states, as they preserve their norm.

A Hermitian operator $\hat{Q}$ that commutes with the Hamiltonian defines a conserved quantity: as you have learned, the expectation value of $\hat{Q}$ is constant in time. We also say that the operator generates a symmetry transformation. Indeed, the operator $e^{i\alpha\hat{Q}}$, with α a real constant that makes the exponent unit-free, is a unitary operator that commutes with the Hamiltonian. As a result, when acting on a nondegenerate energy eigenstate, it preserves the state, and when acting on a degenerate subspace of eigenvectors, it preserves the subspace.

□

Assume the vector space V is finite-dimensional and has an orthonormal basis $(e_1, \ldots, e_n)$. Consider another set of vectors $(f_1, \ldots, f_n)$ where the f's are obtained from the e's by the action of a unitary operator U:

$$f_i = U e_i. \tag{14.5.25}$$

This also means that $e_i = U^\dagger f_i$. The new vectors are also orthonormal:

$$\langle f_i, f_j \rangle = \langle Ue_i, Ue_j \rangle = \langle e_i, e_j \rangle = \delta_{ij}. \tag{14.5.26}$$

They are linearly independent because any list of orthonormal vectors is linearly independent. They span V because dim V linearly independent vectors span V. Thus, the f_i's are an orthonormal *basis*. This is an important result: *the action of a unitary operator on an orthonormal basis gives us another orthonormal basis.*

Let the matrix elements of U in the e-basis be denoted as $U_{ki} = \langle e_k, Ue_i \rangle$. The matrix elements U_{ki}' of U in the f-basis are in fact the same:

$$U'_{ki} = \langle f_k, Uf_i \rangle = \langle Ue_k, Uf_i \rangle = \langle e_k, f_i \rangle = \langle e_k, Ue_i \rangle = U_{ki}. \tag{14.5.27}$$

We first saw this equality when we studied general changes of bases (section 13.5).