## 13.6 EigenvaluesandEigenvectors

In quantum mechanics we need to consider the eigenvalues and eigenstates of Hermitian operators acting on complex vector spaces. These operators are called observables, and their eigenvalues represent possible results of a measurement. In order to acquire a better perspective on these matters, we consider the eigenvalue/eigenvector problem more generally.

One way to understand the action of an operator $T \in \mathcal{L}(V)$ on a vector space $V$ is to describe how it acts on subspaces of $V$. Let $U$ denote a subspace of $V$. In general, the action of $T$ may take elements of $U$ outside $U$. We have a noteworthy situation if $T$ acting on any element of $U$ gives an element of $U$. In this case $U$ is said to be **invariant under** $T$, and $T$ is then a well-defined linear operator on $U$. A very interesting situation arises if a suitable list of invariant subspaces builds up the space $V$ as a direct sum.

Of all subspaces, one-dimensional subspaces are the simplest. Given some nonzero vector $u \in V$, one can consider the one-dimensional subspace $U$ spanned by $u$:

$$U = \{cu \colon c \in \mathbb{F}\}. \tag{13.6.1}$$

We can ask if the one-dimensional subspace $U$ is invariant under $T$. For this $Tu$ must be equal to a number times $u$, as this guarantees that $Tu \in U$. Calling the number $\lambda$, we write

$$T u = \lambda u. \tag{13.6.2}$$

This equation is so ubiquitous that names have been invented to label the objects involved. The number $\lambda \in \mathbb{F}$ is called an **eigenvalue** of the linear operator $T$ *if there is a nonzero vector $u \in V$* such that the equation above is satisfied. It is convenient to call any vector that satisfies (13.6.2) for a given $\lambda$ an **eigenvector** of $T$ corresponding to $\lambda$. In doing so we are including the zero vector as a solution and thus as an eigenvector. Note that with these definitions, having an eigenvalue *means* having associated eigenvectors.

Suppose we find for some specific $\lambda$ a nonzero vector $u$ satisfying (13.6.2). Then it follows that $cu$, for any $c \in \mathbb{F}$, also satisfies the equation so that the solution space of the equation includes the subspace $U$ spanned by $u$, which is an invariant subspace under $T$.

It can often happen that for a given $\lambda$ there are several linearly independent eigenvectors. In this case we say that the eigenvalue $\lambda$ is **degenerate**. The full invariant subspace associated with a degenerate eigenvalue is higher dimensional, and it is spanned by a maximal set of linearly independent eigenvectors; the dimension of this space is called the **geometric multiplicity** of the eigenvalue. The set of eigenvalues of $T$ is called the **spectrum** of $T$.

Equation (13.6.2) is equivalent to

$$(T - \lambda \mathbb{1})\, u = 0, \tag{13.6.3}$$

for some nonzero $u$, so that $T - \lambda \mathbb{1}$ has a nonzero null space and is therefore not injective and not invertible:

$$\boxed{\lambda \text{ is an eigenvalue } \Longleftrightarrow (T - \lambda \mathbb{1}) \text{ is not injective nor invertible.}} \tag{13.6.4}$$

We also note that

$$\boxed{\text{eigenvectors of } T \text{ with eigenvalue } \lambda = \text{null}\,(T - \lambda \mathbb{1}).} \tag{13.6.5}$$

The null space of $T$ is simply the subspace of eigenvectors of $T$ with eigenvalue zero.

It should be noted that the eigenvalues of $T$ and the associated invariant subspaces of eigenvectors are basis independent objects: nowhere in our discussion did we have to invoke the use of a basis. Below, we will review the familiar calculation of eigenvalues and eigenvectors using a matrix representation of the operator $T$.

**Example 13.19.** *Rotation operator in three dimensions.*
Take a real three-dimensional vector space $V$ (our space to great accuracy!). Consider the rotation operator $T$ that rotates all vectors by a fixed angle about the $z$-axis. To find eigenvalues and eigenvectors, we just think of the invariant subspaces. We must ask: Which vectors do not change their direction as a result of this rotation? Only the vectors along the $z$-direction satisfy this condition. So the vector space spanned by $\mathbf{e}_z$ is the invariant subspace, or the space of eigenvectors. The eigenvectors are associated with the eigenvalue $\lambda = 1$ since the vectors are not scaled by the rotation.

□

**Example 13.20.** *Rotation operator in two dimensions.*
Now consider the case where $T$ is a rotation by ninety degrees on a two-dimensional *real* vector space $V$. Are there one-dimensional subspaces invariant under $T$? No, *all* vectors are rotated; none is left invariant. Thus, there are *no eigenvalues* or, of course, eigenvectors. If you tried calculating the eigenvalues by the usual recipe, you would find complex numbers. A complex eigenvalue is meaningless in a real vector space.

□

Although we will not prove the following result, it follows from the facts we have introduced and no extra machinery. It is of interest, being completely general and valid for both real and complex vector spaces:

**Theorem 13.6.1.** *Let $T \in \mathcal{L}(V)$, and assume $\lambda_1, \ldots, \lambda_n$ are distinct eigenvalues of $T$ and $u_1, \ldots, u_n$ are corresponding nonzero eigenvectors. Then $(u_1, \ldots, u_n)$ are linearly independent.*

*Comments:* Note that we cannot ask whether the eigenvectors are orthogonal to each other as we have not yet introduced an inner product on the vector space $V$. There may be more than one linearly independent eigenvector associated with some eigenvalues. In that case any one eigenvector will do. Since an $n$-dimensional vector space $V$ does not have more than $n$ linearly independent vectors, the theorem implies that no linear operator on $V$ can have more than $n$ distinct eigenvalues.

We saw that some linear operators in real vector spaces can fail to have eigenvalues. Complex vector spaces are nicer:

**Theorem 13.6.2.** *Every linear operator on a finite-dimensional complex vector space has at least one eigenvalue.*

*Proof.* This claim means there is at least a one-dimensional invariant subspace spanned by a nonzero eigenvector. The above theorem is a fundamental result, provable with simple tools. The key idea is to consider an arbitrary nonzero vector $v$ in the $n$-dimensional vector space $V$ and to build the list $(v, Tv, T^2v, \ldots, T^nv)$ of $n+1$ vectors. If some entry different from the first vanishes, it would mean that zero is an eigenvalue of $T$ and that the claim holds. Indeed, if $T^kv = 0$ is the first term in the list that vanishes, it means $T^{k-1}v$ is a $T$ eigenvector with eigenvalue zero. Assume now that none of the vectors in the list vanishes. In that case, with the list longer than the dimension of $V$, the vectors are not linearly dependent, and there is some linear relation of the form

$$a_0v + a_1Tv + \cdots + a_nT^nv = 0, \tag{13.6.6}$$

which is satisfied with some set of coefficients $a_i \in \mathbb{C}$, not all of them zero. Let $m$ be the highest value for which $a_m \neq 0$ in this relation. It follows that

$$a_0 v + a_1 T v + \cdots + a_m T^m v = 0, \tag{13.6.7}$$

with $m \leq n$ but also $m \geq 1$ since $v \neq 0$. This equation can be written as

$$(a_0 \mathbb{1} + a_1 T + \cdots + a_m T^m) v = 0. \tag{13.6.8}$$

Inspired by this expression, consider the polynomial $p(z)$ defined by

$$p(z) = a_0 + a_1 z + \cdots a_m z^m. \tag{13.6.9}$$

By the fundamental theorem of algebra, any degree $m$ polynomial over the complex numbers can be written as the product of $m$ linear factors:

$$p(z) = a_0 + a_1 z + \cdots a_m z^m = a_m (z - \lambda_1) \cdots (z - \lambda_m), \tag{13.6.10}$$

where the various $\lambda_i$ may include repeated values. It follows that the above factorization applies to (13.6.8) and allows us to conclude that

$$a_m (T - \lambda_1 \mathbb{1}) \cdots (T - \lambda_m \mathbb{1})\, v = 0. \tag{13.6.11}$$

If all the $(T - \lambda_i \mathbb{1})$ factors above were injective, the equality could not hold: the left-hand side would be nonzero. Therefore, at least one factor $(T - \lambda_k \mathbb{1})$ must fail to be injective. This shows $\lambda_k$ is an eigenvalue of $T$.

□

In order to efficiently find eigenvalues and eigenvectors, one usually considers determinants. When $\lambda$ is an eigenvalue, $T - \lambda \mathbb{1}$ is not invertible, and in any basis, the matrix representative of $T - \lambda \mathbb{1}$ is noninvertible. A matrix is noninvertible if and only if it has zero determinant, therefore,

$$\boxed{\lambda \ \text{ is an eigenvalue} \iff \det(T - \lambda \mathbb{1}) = 0.} \tag{13.6.12}$$

In an $n$-dimensional vector space, the condition for $\lambda$ to be an eigenvalue of $T$ is

$$\det \begin{pmatrix} T_{11} - \lambda & T_{12} & \dots & T_{1n} \\ T_{21} & T_{22} - \lambda & \dots & T_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ T_{n1} & T_{n2} & \dots & T_{nn} - \lambda \end{pmatrix} = 0. \tag{13.6.13}$$

The left-hand side, when computed and expanded out, is a polynomial $f(\lambda)$ in $\lambda$ of degree $n$ called the *characteristic polynomial*:

$$f(\lambda) = \det(T - \lambda\mathbb{1}) = (-\lambda)^n + b_{n-1}\lambda^{n-1} + \cdots + b_1\lambda + b_0, \tag{13.6.14}$$

where the $b_i$ are constants calculable in terms of the $T_{ij}$'s. The equation

$$f(\lambda) = 0 \tag{13.6.15}$$

determines all eigenvalues. Over the complex numbers, the characteristic polynomial can be factorized, again, by the fundamental theorem of algebra,

$$f(\lambda) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n). \tag{13.6.16}$$

The notation does not preclude the possibility that some of the $\lambda_i$'s may be equal. The $\lambda_i$'s are the eigenvalues, since they lead to $f(\lambda) = 0$ for $\lambda = \lambda_i$. Even when all factors in the characteristic polynomial are identical, we still have one eigenvalue. For any eigenvalue $\lambda$, the operator $(T - \lambda\mathbb{1})$ is not injective and thus has a nonvanishing null space. Any vector in this null space is an eigenvector with eigenvalue $\lambda$.

If all eigenvalues of $T$ are different, the spectrum of $T$ is said to be **nondegenerate**. If an eigenvalue $\lambda_i$ appears $k$ times, the characteristic polynomial includes the factor $(\lambda - \lambda_i)^k$, and $\lambda_i$ is said to be a degenerate eigenvalue with **algebraic multiplicity** $k$. In general, the geometric multiplicity of an eigenvalue is less than or equal to the algebraic multiplicity, implying that the number of linearly independent eigenvectors with eigenvalue $\lambda_i$ can be less than or equal to $k$. For Hermitian operators in complex vector spaces, however, the two multiplicities are the same.

**Exercise 13.9.** *Show that the matrix* $\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$ *has an eigenvalue with algebraic multiplicity two and geometric multiplicity one.*

**Example 13.21.** *Eigenvectors of the annihilation operator $\hat{a}$ of the harmonic oscillator.*
Assume $\psi_\lambda$ is an eigenstate of $\hat{a}$ with eigenvalue $\lambda$:

$$\hat{a}\,\psi_\lambda = \lambda\psi_\lambda. \tag{13.6.17}$$

We wish to find the allowed values of $\lambda$ and the corresponding expressions for the eigenvectors $\psi_\lambda$. We write a general ansatz for the state $\psi_\lambda$:

$$\psi_\lambda = \varphi_0 + \sum_{n=1}^{\infty} c_n \varphi_n. \tag{13.6.18}$$

In here we used the normalization ambiguity to scale the coefficient of the ground state to one. The ground state must be present—if not, the leading term of the state would be some $\varphi_k$ with $k \geq 1$, and the action of $\hat{a}$ on the state would create a term $\sim \varphi_{k-1}$ not present in $\psi_\lambda$, making it impossible to satisfy the eigenvalue equation. Acting with $\hat{a}$ on the state, we have

$$\hat{a}\psi_\lambda = c_1\varphi_0 + \sum_{n=2}^{\infty} c_n \sqrt{n}\varphi_{n-1} = c_1\Big(\varphi_0 + \sum_{n=1}^{\infty} \frac{c_{n+1}}{c_1}\sqrt{n+1}\varphi_n\Big). \tag{13.6.19}$$

For the right-hand side to be equal to $\lambda\psi_\lambda$, we must have $c_1 = \lambda$. This suggests that, in fact, we may get a solution of the eigenvalue equation for arbitrary $\lambda \in \mathbb{C}$. Writing $c_1 = \lambda$, the above equation becomes

$$\hat{a}\psi_\lambda = \lambda\Big(\varphi_0 + \sum_{n=1}^{\infty} \frac{c_{n+1}}{\lambda}\sqrt{n+1}\varphi_n\Big). \tag{13.6.20}$$

For the state inside the parentheses to be $\psi_\lambda$, as in (13.6.18), we must have

$$\frac{c_{n+1}}{\lambda}\sqrt{n+1} = c_n \quad \Rightarrow \quad c_{n+1} = \frac{\lambda}{\sqrt{n+1}} c_n, \quad n = 1, 2, \ldots . \tag{13.6.21}$$

With $c_1 = \lambda$, the solution for any $n \geq 1$ is quickly checked to be

$$c_n = \frac{\lambda^n}{\sqrt{n!}}, \quad n \geq 1. \tag{13.6.22}$$

We can then rewrite the state $\psi_\lambda$ in (13.6.18):

$$\psi_\lambda = \varphi_0 + \sum_{n=1}^{\infty} \frac{\lambda^n}{\sqrt{n!}} \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}}\varphi_0 = \Big(1 + \sum_{n=1}^{\infty} \frac{(\lambda\hat{a}^\dagger)^n}{n!}\Big)\varphi_0. \tag{13.6.23}$$

The factor in parentheses is in fact the expansion of an exponential. We can then write

$$\hat{a}\,\psi_\lambda = \lambda\psi_\lambda, \qquad \psi_\lambda = \exp(\lambda\,\hat{a}^\dagger)\,\varphi_0, \quad \lambda \in \mathbb{C}. \tag{13.6.24}$$

There is no quantization of the spectrum: all complex values of $\lambda$ are allowed. The eigenvalues can be complex because $\hat{a}$ is not Hermitian. The states $\psi_\lambda$, called coherent states, were considered in problem 9.9 and will be studied in more detail in chapter 17. The ground state $\varphi_0$ is a coherent state $\psi_0$ with $\lambda = 0$.

□

Theorem 13.6.2, establishing that any linear operator on a finite-dimensional complex vector space has an eigenvalue, does not apply for $\hat{a}$ because the relevant vector space is infinite-dimensional. We were not guaranteed to find an eigenvalue, but in fact, we found an infinite number of them. This need not always be the case:

**Exercise13.10.** *Explain why the creation operator $\hat{a}^\dagger$ has no eigenvalues in the harmonic oscillator state space.*