13
Vector Spaces and Operators

We define real and complex vector spaces. Using subspaces, we introduce the concept of the direct sum representation of a vector space. We define spanning sets, linearly independent sets, basis vectors, and the concept of dimensionality. Linear operators are linear functions from a vector space to itself. The set of linear operators on a vector space itself forms a vector space, with a product structure that arises by composition. While the product is associative, it is not in general commutative. We discuss when operators are injective, surjective, and invertible. We explain how operators can be represented by matrices once we have made a choice of basis vectors on the vector space. We derive the relation between matrix representations in different bases and show that traces and determinants are basis independent. Finally, we discuss eigenvectors and eigenvalues of operators, connecting them to the idea of invariant subspaces.

13.1 Vector Spaces

In quantum mechanics the state of a physical system is a *vector* in a *complex vector space*. Observables are linear operators—in fact, Hermitian operators acting on this complex vector space. The purpose of this chapter is to learn the basics of vector spaces and the operators that act on them.

Complex vector spaces are somewhat different from the more familiar real vector spaces. They have more powerful properties. In order to understand complex vector spaces, it is useful to compare them often to their real-dimensional friends. In a vector space, one has vectors and numbers, also sometimes referred to as scalars. We can add vectors to get vectors, and we can multiply vectors by numbers to get vectors. If the numbers we use are real, we have a real vector space. If the numbers we use are complex, we have a complex vector space. More generally, the numbers we use belong to what is called a *field*, denoted by the letter $\mathbb{F}$. We will discuss just two cases, $\mathbb{F} = \mathbb{R}$, meaning the numbers are real, and $\mathbb{F} = \mathbb{C}$, meaning the numbers are complex.

The definition of a vector space is the same for any field $\mathbb{F}$. A vector space V is a set of vectors with an operation of **addition** (+) that assigns an element $u + v \in V$ to each $u, v \in V$. Because the operation of addition takes two elements of V to another element of V, we say that V is *closed* under addition. There is also a **scalar multiplication** by elements of $\mathbb{F}$, with $av \in V$ for any $a \in \mathbb{F}$ and $v \in V$. This means the space V is also closed under multiplication by numbers. These operations must satisfy the following additional properties:

1. $u + v = v + u \in V$ for all $u, v \in V$ (addition is commutative).
2. $u + (v + w) = (u + v) + w$, and $(ab)u = a(bu)$ for any $u, v, w \in V$ and $a, b \in \mathbb{F}$ (associativity of addition and scalar multiplication).
3. There is a vector $0 \in V$ such that $0 + u = u$ for all $u \in V$ (additive identity).
4. For each $v \in V$, there is a $u \in V$ such that $v+u = 0$ (additive inverse).
5. The element $1 \in \mathbb{F}$ satisfies $1v = v$ for all $v \in V$ (multiplicative identity).
6. $a(u + v) = au + av$, and $(a + b)v = av + bv$ for every $u, v \in V$ and $a, b \in \mathbb{F}$ (distributive property).

This definition is very efficient. Several familiar properties follow from it by short proofs that we will not give but that are not complicated and you may try to produce. For example:

- The additive identity is unique: any vector $0'$ that acts like 0 is actually equal to 0.
- $0v = 0$, for any $v \in V$, where the first zero is a number, and the second one is a vector. This means that the number zero acts as expected when multiplying a vector.

- $a0 = 0$, for any $a \in \mathbb{F}$. Here both zeroes are vectors. This means that the zero vector multiplied by any number is still the zero vector.
- The additive inverse of any vector $v \in V$ is unique. It is denoted by $-v$ and in fact $-v = (-1)v$.

We must stress that while the numbers in $\mathbb{F}$ are sometimes real or complex, we *do not* speak of the vectors themselves as real or complex. A vector multiplied by a complex number, for example, is not said to be a complex vector. The vectors in a real vector space are not said to be real, and the vectors in a complex vector space are not said to be complex.

The definition of a vector space does *not* introduce a multiplication of vectors. Only in very special cases is there a natural way to multiply vectors to give vectors. One such example is the cross product in three spatial dimensions.

As is customary in the mathematics literature, vectors are denoted by symbols such as v, u, w, without arrows added on top or the use of boldface. In physics we sometimes use arrows or boldface. We hope the various notations will not cause confusion.

A guide to the examples In this and the following chapter, as we introduce the mathematical concepts we illustrate their use with examples from quantum systems we have examined before. This should help the reader appreciate the material. Moreover, in this way the reader will review those important quantum systems. For the case of spin one-half systems, the examples develop considerably the theory discussed in chapter 12. The examples also discuss the harmonic oscillator. Finally, the examples include mathematical techniques used in computations, such as index notation and commutator algebra. Below we list the examples so the reader can access them quickly. Examples of a purely mathematical nature are not included in the lists below.

Examples and developments for spin one-half systems:

- Example 13.2. Vector space $\mathbb{C}^2$ for spin one-half.
- Example 13.7. The real vector space of 2×2 Hermitian matrices.
- Example 13.11. Working with Pauli matrices.

- Example 13.12. Is there a linear operator that reverses the direction of all spin states?
- Example 13.14. Null space and range of spin operators on $\mathbb{C}^2$.
- Example 13.22. Exponentials of linear combinations of Pauli matrices.
- Example 14.1. Inner product in $\mathbb{C}^2$ for spin one-half.
- Example 14.2. Inner product of spin states.
- Example 14.4. An orthonormal basis of Hermitian matrices in two dimensions.
- Example 14.7. Orthogonal projector onto a spin state $|\mathbf{n}\rangle$.
- Section 14.7. Rotation operators for spin states.
- Example 14.13. Orthogonal projector onto a spin state $|\mathbf{n}\rangle$ revisited.

Examples and developments for the simple harmonic oscillator:

- Example 13.4. Fock states of the one-dimensional harmonic oscillator.
- Example 13.8. State space $\mathcal{H}$ of the simple harmonic oscillator.
- Example 13.16. Right inverse for the annihilation operator $\hat{a}$.
- Example 13.18. Matrix representation for the harmonic oscillator $\hat{a}$ and $\hat{a}^\dagger$ operators.
- Example 13.21. Eigenvectors of the annihilation operator $\hat{a}$ of the harmonic oscillator.
- Example 14.16. Harmonic oscillator in bra-ket notation.

Examples for computational techniques:

- Example 13.10. Review of index manipulations.
- Section 13.7. Functions of linear operators and key identities.
- Example 13.23. Five ways to do a computation.
- Example 14.3. Inner product for operators.
- Example 14.12. Trace of the operator $|u\rangle\langle w|$.
- Example 14.14. Adjoint in bra-ket description of operators.

Additional examples:

- Example 13.3. Vector space $\mathbb{C}^{2\ell+1}$ for orbital angular momentum ℓ.
- Example 13.5. The state space of the infinite square well.
- Example 13.9. State space of hydrogen atom bound states.

- Example 14.11. From Hermitian to unitary operators.

Example 13.1. *A few vector spaces.*

1. The vector space $\mathbb{C}^N$ is the set of N-component vectors:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix}, \quad a_i \in \mathbb{C}, \quad i = 1, 2, \dots, N. \tag{13.1.1}$$

   We add two vectors by adding the corresponding components:

$$\begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_N \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_N + b_N \end{pmatrix}. \tag{13.1.2}$$

   Multiplication by a number $\alpha \in \mathbb{C}$ is defined by multiplying each component by α:

$$\alpha \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} = \begin{pmatrix} \alpha\, a_1 \\ \vdots \\ \alpha\, a_N \end{pmatrix}. \tag{13.1.3}$$

   The zero vector is the vector with the number zero for each entry. You should verify that all the axioms are then satisfied. This is the classic example of a complex vector space. If all entries and numbers are real, this would be a real vector space.

2. Here is a slightly more unusual example in which matrices are the vectors of the vector space. Consider the set of $M \times N$ matrices with complex entries:

$$\begin{pmatrix} a_{11} & \dots & a_{1N} \\ \vdots & & \vdots \\ a_{M1} & \dots & a_{MN} \end{pmatrix}, \quad a_{ij} \in \mathbb{C}. \tag{13.1.4}$$

   We define addition as follows:

$$\begin{pmatrix} a_{11} & \dots & a_{1N} \\ \vdots & & \vdots \\ a_{M1} & \dots & a_{MN} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1N} \\ \vdots & & \vdots \\ b_{M1} & \dots & b_{MN} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1N} + b_{1N} \\ \vdots & & \vdots \\ a_{M1} + b_{M1} & \dots & a_{MN} + b_{MN} \end{pmatrix}. \tag{13.1.5}$$

Multiplication by a constant $f \in \mathbb{C}$ ends up multiplying all entries:

$$f\begin{pmatrix} a_{11} & \dots & a_{1N} \\ \vdots & & \vdots \\ a_{M1} & \dots & a_{MN} \end{pmatrix} = \begin{pmatrix} fa_{11} & \dots & fa_{1N} \\ \vdots & & \vdots \\ fa_{M1} & \dots & fa_{MN} \end{pmatrix}. \tag{13.1.6}$$

The zero matrix is defined as the matrix in which all entries are the number zero. With these definitions the set of $M \times N$ matrices forms a complex vector space.

3. Consider the set of $N \times N$ Hermitian matrices. These are matrices with complex entries that are left invariant by the successive operations of transposition and complex conjugation. Curiously, the set of $N \times N$ Hermitian matrices form a *real* vector space. This is because multiplication by real numbers preserves the property of Hermiticity, while multiplication by complex numbers does not. This illustrates the earlier claim that we should not use the labels *real* or *complex* for the vectors themselves.
4. Here is another slightly unusual example in which polynomials are the vectors. Consider the set □(□) of polynomials. A polynomial $p \in$□($\mathbb{F}$) is a function from $\mathbb{F}$ to $\mathbb{F}$: acting on the variable $z \in \mathbb{F}$, it gives a value $p(z) \in \mathbb{F}$. Each nonzero polynomial p is defined by coefficients a_0, a_1, $\dots a_n \in \mathbb{F}$, with n a finite, nonnegative integer called the degree of the polynomial:

$$p(z) = a_0 + a_1 z + a_2 z^2 + \dots + a_n z^n, \quad a_n \neq 0. \tag{13.1.7}$$

The zero polynomial $p(z) = 0$ has $n = 0$ and $a_0 = 0$. The addition of polynomials works as expected. If p_1, $p_2 \in$□($\mathbb{F}$), then $p_1 + p_2 \in$□($\mathbb{F}$) is defined by

$$(p_1 + p_2)(z) = p_1(z) + p_2(z), \tag{13.1.8}$$

and multiplication works as $(ap)(z) = ap(z)$, for $a \in \mathbb{F}$ and $p \in$□($\mathbb{F}$). The zero vector is the zero polynomial. The space □(□) of all polynomials forms a vector space over $\mathbb{F}$.

This vector space has a simple generalization when z is just a formal variable not valued in $\mathbb{F}$, while the coefficients a_k are still valued in $\mathbb{F}$.

This vector space actually represents an interesting subspace of states of the quantum harmonic oscillator (example 13.4).

5. Consider the set $\mathbb{F}^\infty$ of infinite sequences $(x_1, x_2, \ldots)$ of elements $x_i \in \mathbb{F}$. Here

$$\begin{aligned}(x_1,x_2,\ldots)+(y_1,y_2,\ldots) &= (x_1+y_1,x_2+y_2,\ldots),\\ a(x_1,x_2,\ldots) &= (ax_1,ax_2,\ldots), \quad a\in\mathbb{F}.\end{aligned} \tag{13.1.9}$$

This is a vector space over $\mathbb{F}$.

6. The set of complex functions $f(x)$ on an interval $x \in [0, L]$ form a vector space over $\mathbb{C}$. Here the functions are the vectors of the vector space. The required definitions are

$$(f_1+f_2)(x) = f_1(x)+f_2(x), \quad (af)(x) = af(x) \tag{13.1.10}$$

with $f_1(x)$ and $f_2(x)$ complex valued functions on the interval and a a complex number. This vector space contains, for example, the wave functions of a particle in a one-dimensional potential that confines it to the interval $x \in [0, L]$.

□

Let us now see in more detail how some of the above vector spaces are suitable to the description of familiar quantum systems.

Example 13.2. *Vector space $\mathbb{C}^2$ for spin one-half.*
We had our first look at spin one-half in chapter 12. The state space there was that of two-component complex vectors (as in item 1 of example 13.1). The quantum states Ψ of the spin one-half particle take the form

$$\Psi = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \text{ with } c_1,c_2\in\mathbb{C}. \tag{13.1.11}$$

It takes just two complex numbers to specify completely the spin state of the particle. Note how this differs from the data required to specify the position state of the particle: a whole wave function $\psi(x)$ worth of data. Using the rule for multiplying vectors by complex constants, we can rewrite Ψ as

$$\Psi = c_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{13.1.12}$$

The column vectors with a single 1 and a single 0 have been given names:

$$|z;+\rangle = |1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |z;-\rangle = |2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{13.1.13}$$

We identified $|z;+\rangle$ as the state of a particle with its spin pointing in the positive z-direction and $|z;-\rangle$ as the state of a particle with its spin pointing in the negative z-direction. With this notation the state in (13.1.11) is $\Psi = c_1|z;+\rangle + c_2|z;-\rangle$.

□

Example 13.3. *Vector space $\mathbb{C}^{2\ell+1}$ for orbital angular momentum ℓ.*
We learned in chapter 10 that in a central potential energy eigenstates are described with wave functions of the form $\psi(r, \theta, \phi) = R(r)F(\theta, \phi)$. The angular dependence $F(\theta, \phi)$ determines the angular momentum of the state. Restricting ourselves to quantum states of angular momentum ℓ means that we pick wave functions $F_\ell(\theta, \phi)$ that are $\hat{L}^2$ eigenstates:

$$\hat{\mathbf{L}}^2 F_\ell(\theta, \phi) = \hbar^2 \ell(\ell+1) F_\ell(\theta, \phi). \tag{13.1.14}$$

In fact, we found that the wave functions $Y_{\ell m}(\theta, \phi)$ with $m = -\ell, \ldots, \ell$ all have angular momentum ℓ. Indeed, from (10.5.32) we have that

$$\begin{aligned} \hat{\mathbf{L}}^2 \, Y_{\ell m} &= \hbar^2 \, \ell(\ell+1) \, Y_{\ell m}, \\ \hat{L}_z \, Y_{\ell m} &= \hbar m \, Y_{\ell m}, \qquad m = -\ell, \ldots, \ell. \end{aligned} \tag{13.1.15}$$

Any angular wave function with angular momentum ℓ must be a linear superposition of the $Y_{\ell m}$'s with various values of m. It is thus natural to define the vector space $\mathbb{C}^{2\ell+1}$ of angular wave functions by encoding general wave functions of angular momentum ℓ into a column vector of size $2\ell + 1$ with arbitrary complex entries:

$$F_\ell(\theta, \phi) = c_1 Y_{\ell\ell} + c_2 Y_{\ell,\ell-1} + \cdots c_{2\ell+1} Y_{\ell,-\ell} \quad \Longleftrightarrow \quad \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{2\ell+1} \end{pmatrix}, \quad c_i \in \mathbb{C}. \tag{13.1.16}$$

The addition of vectors is natural, and so is multiplication by a constant. The spherical harmonics themselves can be thought of as simple vectors with a single nonvanishing entry equal to one:

$$Y_{\ell\ell} = e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots \quad , \; Y_{\ell,-\ell} = e_{2\ell+1} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \tag{13.1.17}$$

For each possible value of $\ell = 0, 1, \ldots$, there is a state space $\mathbb{C}^{2\ell+1}$ of states. For $\ell = 0$, the space is just the space of complex constants $\mathbb{C}$.

□

Example 13.4. *Fock states of the one-dimensional harmonic oscillator.* There exist interesting states, called *Fock states*, that consist of *finite* linear superpositions of energy eigenstate states $\varphi_n \sim (\hat{a}^\dagger)^n\varphi_0$, with φ_0 the ground state (see chapter 9). A Fock state ψ_k takes the form

$$\psi_k \;=\; \big(\gamma_0 + \gamma_1\hat{a}^\dagger + \gamma_2(\hat{a}^\dagger)^2 + \cdots + \gamma_k(\hat{a}^\dagger)^k\big)\varphi_0, \quad \gamma_i \in \mathbb{C}. \tag{13.1.18}$$

Fock states are in one-to-one correspondence with the vector space of *formal* polynomials discussed in example 13.1, item 4. We consider the space □(□) of polynomials with complex coefficients and a formal variable z identified with $\hat{a}^\dagger$. In this correspondence, the above state ψ_k is unambiguously associated with the polynomial $p_k(z)$ below:

$$p_k(z) = \gamma_0 + \gamma_1 z + \gamma_2 z^2 + \cdots + \gamma_k z^k. \tag{13.1.19}$$

Fock states are manifestly normalizable; they are, after all, a finite sum of energy eigenstates.

Not all states in the state space of the harmonic oscillator are Fock states. We will find normalizable coherent states and squeezed states that are not polynomials in the creation operator acting on the ground state. They are, instead, exponentials of linear and quadratic functions of the creation operator. The normalizability of non-Fock states must be checked since they are built as *infinite* linear superpositions of energy eigenstates.

□

Example 13.5. *The state space of the infinite square well.*

The energy eigenstates of the infinite square well $x \in [0, a]$ (section 6.4) are represented by orthonormal wave functions:

$$\psi_n(x) = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right) \quad \text{with} \quad n = 1, 2, \cdots . \tag{13.1.20}$$

The specification of a state of the square well is tantamount to the specification of an infinite sequence $(c_1, c_2, c_3, \ldots)$ of complex numbers that result in the state $\psi(x)$:

$$\psi(x) = \sum_{n=1}^{\infty} c_n \psi_n(x) = c_1 \psi_1(x) + c_2 \psi_2(x) + \cdots . \tag{13.1.21}$$

This is in fact the space $\mathbb{C}^\infty$ of infinite sequences presented in example 13.1, item 5. The states of the square well, however, must be normalizable so that $\|\psi\|^2 = \sum_{n=1}^{\infty} |c_n|^2 < \infty$. The infinite sequences in $\mathbb{C}^\infty$ satisfying this constraint still form a vector space. Item 6 of that same example examines the space of complex functions on an interval. This is the space we are considering when restricted to functions that vanish at the end points and are square integrable.