13.3 LinearOperators

A linear map is a particular kind of function from one vector space V to another vector space W. When the linear map takes the vector space V to itself, we call the linear map a linear operator. We will focus our attention on these operators. In quantum mechanics linear operators produce the time evolution of states. Moreover, physical observables are associated with linear operators.

A **linearoperator** T on a vector space V is a function that takes V to V with the following properties:

1. $T(u + v) = Tu + Tv$, for all $u, v \in V$.

2. $T(au) = aTu$, for all $a \in \mathbb{F}$ and $u \in V$.

In the above notation, Tu, for example, means the result of the action of the operator T on the vector u. It could also be written as $T(u)$, but it is simpler to write it as Tu, in a way that makes the action of T on u look "multiplicative."

A simple consequence of the axioms is that the action of a linear operator on the zero vector is the zero vector:

$$T0 = 0. \tag{13.3.1}$$

This follows from $Tu = T(u + 0) = Tu + T\,0$ and canceling the common Tu term.

Let us consider a few examples of linear operators;

1. Let $V = □[x]$ denote the space of real polynomials $p(x)$ of a real variable x with real coefficients. Here are two linear operators T and S on V:
   - Let T denote differentiation: $Tp = p'$ where $p' \equiv \frac{dp}{dx}$. This operator is linear because

$$\begin{aligned} T(p_1 + p_2) = {} & (p_1 + p_2)' = p_1' + p_2' = Tp_1 + Tp_2, \\ T(ap) = {} & (ap)' = ap' = a\,Tp. \end{aligned} \tag{13.3.2}$$

   - Let S denote multiplication by x: $Sp = xp$. S is also a linear operator.

2. In the space $\mathbb{F}^\infty$ of infinite sequences, define the **left-shift** operator L by

$$L(x_1, x_2, x_3, \ldots) = (x_2, x_3, \ldots). \tag{13.3.3}$$

By shifting to the left, we lose the information about the first entry, but that is perfectly consistent with linearity. We also have the **right-shift** operator R that acts by shifting to the right and creating a new first entry as follows:

$$R(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots). \tag{13.3.4}$$

The first entry after the action of R is zero. It could not be any other number because the zero element (a sequence of all zeroes) should be mapped to itself (by linearity).

3. For any vector space V, we define the **zero operator** 0 that, acting on any vector in V, maps it to the zero vector: $0v = 0$ for all $v \in V$. This map is very simple, almost trivial, but certainly linear. Note that now we have the zero number, the zero vector, and the zero operator, all denoted by the symbol 0.
4. For any vector space V, we define the **identity operator** $\mathbb{1}$ that leaves all vectors in V invariant: $\mathbb{1}v = v$ for all $v \in V$.

On any vector space V, there are many linear operators. We call $\mathcal{L}(V)$ the set of all linear operators on V. Since operators on V can be added and can also be multiplied by numbers, the set $\mathcal{L}(V)$ **is itself a vector space**, where the vectors are the operators. Indeed, for any two operators $S, T \in \mathcal{L}(V)$ we have the natural definition

$$\begin{aligned} (S+T)v &= Sv + Tv, \\ (aS)v &= a(Sv). \end{aligned} \tag{13.3.5}$$

A vector space must have an additive identity. Here it is an operator that can be added to other operators with no effect. The additive identity in the vector space $\mathcal{L}(V)$ is the zero operator on V, considered in (3) above.

In the vector space $\mathcal{L}(V)$, there is a surprising new structure: the vectors (the operators!) can be naturally multiplied. There is a **multiplication of linear operators** that gives a linear operator: we just let one operator act first and the other next! So given $S, T \in \mathcal{L}(V)$, we define the operator ST as

$$(ST)v \equiv S(Tv). \tag{13.3.6}$$

We easily verify linearity:

$$(ST)(u+v) = S(T(u+v)) = S(Tu+Tv) = S(Tu) + S(Tv) = (ST)(u) + (ST)(v), \tag{13.3.7}$$

and you can also verify that $(ST)(av) = a(ST)(v)$.

The product just introduced in the space of linear operators is **associative**. This is a fundamental property of operators and means that for S, T, U, linear operators

$$S(TU) = (ST)U. \tag{13.3.8}$$

This equality holds because acting on any vector v both the left-hand side and the right-hand side give $S(T(U(v)))$. The product has an identity element: the identity operator $\mathbb{1}$ of (4). If we have a product, we can ask if the elements (the operators) have inverses. As we will see later, some operators have inverses and some do not.

Finally, and crucially, this product is in general **noncommutative**. We can check this using the two operators T and S of (1), acting on the polynomial $p = x^n$. Since T differentiates and S multiplies by x, we get

$$\begin{aligned}(TS)x^n &= T(Sx^n) = T(x^{n+1}) = (n+1)x^n, \\ (ST)x^n &= S(Tx^n) = S(nx^{n-1}) = nx^n.\end{aligned} \tag{13.3.9}$$

We quantify the failure of commutativity by the difference $TS - ST$, which is itself a linear operator:

$$(TS - ST)x^n = (n+1)x^n - nx^n = x^n = \mathbb{1}\, x^n, \tag{13.3.10}$$

where we inserted the identity operator at the last step. Since this relation is true acting on x^n, for any $n \geq 0$, it holds by linearity acting on any polynomial—namely, on any element of the vector space. So we can simply write

$$[\,T, S\,] = \mathbb{1}, \tag{13.3.11}$$

where we introduced the **commutator** $[\cdot, \cdot]$ of two linear operators X, Y, defined by

$$[X, Y] \equiv XY - YX. \tag{13.3.12}$$

Exercise 13.4. *Calculate the commutator* $[L, R]$ *of the left-shift and right-shift operators. Express your answer using the identity operator and the operator* P_1 *defined by* $P_1(x_1, x_2, \ldots) = (x_1, 0, 0, \ldots)$.

Example 13.11. *Working with Pauli matrices.*

The Pauli matrices σ_i, with $i = 1, 2, 3$, or the associated spin operators $S_i = \frac{\hbar}{2}\sigma_i$ are indeed operators on $\mathbb{C}^2$, the vector space of spin states (see example 13.2). We should be able to manipulate these 2×2 matrices efficiently. We first recall their explicit form

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{13.3.13}$$

These matrices are Hermitian; in fact, together with the identity matrix they span the real vector space of 2 × 2 Hermitian matrices (example 13.7). They are also traceless:

$$\mathrm{tr}\,\sigma_i = 0, \quad i = 1, 2, 3. \tag{13.3.14}$$

The Pauli matrices square to the identity matrix, as one can check explicitly:

$$(\sigma_1)^2 = (\sigma_2)^2 = (\sigma_3)^2 = \mathbb{1}. \tag{13.3.15}$$

This property implies that the eigenvalues of each of the Pauli matrices can only be plus or minus one. Indeed, the eigenvalues of a matrix satisfy the algebraic equation that the matrix satisfies. Thus, the eigenvalues must satisfy $\lambda^2 = 1$, showing that $\lambda = \pm 1$ are the only options. Since the sum of eigenvalues equals the trace, which is vanishing, each Pauli matrix has an eigenvalue $+1$ and an eigenvalue -1.

The commutation relations for the spin operators $[\hat{S}_i, \hat{S}_j] = i\hbar\, \epsilon_{ijk}\hat{S}_k$ together with $\hat{S}_i = \frac{\hbar}{2}\sigma_i$ imply that

$$[\sigma_i, \sigma_j] = 2i\,\epsilon_{ijk}\sigma_k. \tag{13.3.16}$$

Make sure never to confuse the imaginary number i with the index i. If you compute a commutator of Pauli matrices by hand, you might notice a curious property. Take the commutator $[\sigma_1, \sigma_2] = 2i\sigma_3$. If you do the matrix multiplications, you find that $\sigma_1\sigma_2 = i\sigma_3$ while $\sigma_2\sigma_1 = -i\sigma_3$. These two products differ by a sign:

$$\sigma_1\sigma_2 = -\sigma_2\sigma_1. \tag{13.3.17}$$

We say that σ_1 and σ_2 *anticommute*: they can be moved across each other at the cost of a sign. Just as we define the commutator of two operators X, Y by $[X, Y] \equiv XY - YX$, we define the **anticommutator**, denoted by curly brackets, by the following:

$$\text{anticommutator:}\ \{X, Y\} \equiv XY + YX. \tag{13.3.18}$$

In this language we have checked that $\{\sigma_1, \sigma_2\} = 0$, and the property $\sigma_1^2 = \mathbb{1}$, for example, can be rewritten as $\{\sigma_1, \sigma_1\} = 2 \cdot \mathbb{1}$. In fact, you can check (by examining the two remaining cases) that any two different Pauli matrices anticommute:

$$\{\sigma_i, \sigma_j\} = 0, \quad \text{for } i \neq j. \tag{13.3.19}$$

We can improve this equation to make it also work when i is equal to j. We claim that

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}\, \mathbb{1}. \tag{13.3.20}$$

Indeed, when $i \neq j$ the right-hand side vanishes, as needed, and when i is equal to j, the right-hand side gives $2 \cdot \mathbb{1}$, which is also needed since the Pauli matrices square to the identity.

The commutator and anticommutator identities for the Pauli matrices can be summarized in a single equation. This is possible because for any two operators X, Y we have

$$XY = \tfrac{1}{2}\{X, Y\} + \tfrac{1}{2}[X, Y], \tag{13.3.21}$$

as you should confirm by expansion. Applied to the product of two Pauli matrices and using our expressions for the commutator and anticommutator, we get

$$\boxed{\sigma_i \sigma_j = \delta_{ij}\, \mathbb{1} + i\epsilon_{ijk}\, \sigma_k.} \tag{13.3.22}$$

Note that $\sigma_k \sigma_{k+1} = i\sigma_{k+2}$, where we use arithmetic modulo 3 in the subscripts ($4 \equiv 1$, $5 \equiv 2$). The equation for $\sigma_i \sigma_j$ can be recast in vector notation if we introduce the "vector" triplet of Pauli matrices:

$$\boldsymbol{\sigma} \equiv (\sigma_1, \sigma_2, \sigma_3). \tag{13.3.23}$$

We can construct a matrix by the dot product of a vector $\mathbf{a} = (a_1, a_2, a_3)$ with the "vector" $\boldsymbol{\sigma}$. Here the components a_i of $\mathbf{a}$ are assumed to be numbers. We define

$$\mathbf{a} \cdot \boldsymbol{\sigma} \equiv a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3 = a_i\sigma_i. \tag{13.3.24}$$

Note that $\mathbf{a} \cdot \boldsymbol{\sigma}$ is just a single 2×2 matrix. The components of $\mathbf{a}$, being numbers, commute with matrices, and this dot product is commutative: $\mathbf{a} \cdot \boldsymbol{\sigma} = \boldsymbol{\sigma} \cdot \mathbf{a}$. To rewrite (13.3.22) we multiply this equation by $a_i b_j$ to get

$$\begin{aligned} a_i\sigma_i\, b_j\sigma_j &= a_i b_j \delta_{ij}\,\mathbb{1} \;+\; i\,(a_i b_j \epsilon_{ijk})\,\sigma_k \\ &= (\mathbf{a}\cdot\mathbf{b})\,\mathbb{1} \;+\; i\,(\mathbf{a}\times\mathbf{b})_k\,\sigma_k \end{aligned} \tag{13.3.25}$$

so that, finally, we get the matrix equation

$$(\mathbf{a}\cdot\boldsymbol{\sigma})(\mathbf{b}\cdot\boldsymbol{\sigma}) = (\mathbf{a}\cdot\mathbf{b})\,\mathbb{1} + i\,(\mathbf{a}\times\mathbf{b})\cdot\boldsymbol{\sigma}. \tag{13.3.26}$$

This equation holds even if the components of $\mathbf{a}$ and $\mathbf{b}$ are operators, provided the operators commute with the Pauli matrices, as is often the case in applications. Indeed, in deriving the above equation we never had to move any a_i across any b_j. As a simple application, we take $\mathbf{b} = \mathbf{a}$, with components of ordinary numbers. We then have $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$ as well as $\mathbf{a} \times \mathbf{a} = 0$, an equation that can fail when $\mathbf{a}$ has operator components. The above equation then gives

$$(\mathbf{a}\cdot\boldsymbol{\sigma})^2 = |\mathbf{a}|^2\,\mathbb{1}. \tag{13.3.27}$$

When $\mathbf{a}$ is a unit vector $\mathbf{n}$, this becomes

$$(\mathbf{n}\cdot\boldsymbol{\sigma})^2 = \mathbb{1}, \quad \mathbf{n}\cdot\mathbf{n} = 1. \tag{13.3.28}$$

Since $\mathbf{n} \cdot \boldsymbol{\sigma}$ is Hermitian (being a superposition of Pauli matrices with real coefficients) and traceless, it follows that $\mathbf{n} \cdot \boldsymbol{\sigma}$, just like any Pauli matrix, has eigenvalues $\pm$ 1. It thus follows that the spin operator $\hat{S}_{\mathbf{n}} = \frac{\hbar}{2}\mathbf{n}\cdot\boldsymbol{\sigma}$ has eigenvalues $\pm\frac{\hbar}{2}$. This was the reason we could think of $\hat{S}_{\mathbf{n}}$ as a spin operator in the direction of $\mathbf{n}$.

□

Example 13.12. *Is there a linear operator that reverses the direction of all spin states?*

A simple way to define a linear operator on a vector space V is to define its action on a set of *basis* vectors of V. Once you know how the operator acts on the basis vectors, you know by linearity how it acts on arbitrary vectors. It is far more delicate to define a linear operator by stating how it acts on *every* vector in V. In that case one must check the consistency of the definition with linearity.

We ask if there is a linear operator T that reverses the direction of all spin states in $\mathbb{C}^2$. If it existed, it must take an arbitrary spin state $|\mathbf{n};+\rangle$ into the state $|\mathbf{n};-\rangle$, up to a constant. Let us test whether this is possible. If T reverses every spin state, it must send $|+\rangle$ to $|-\rangle$ and vice versa. Of course, in general, it can do this up to nonvanishing constants α and β to be determined:

$$T|+\rangle = \alpha|-\rangle, \qquad T|-\rangle = \beta|+\rangle, \quad \alpha, \beta \in \mathbb{C}. \tag{13.3.29}$$

Let us test this on the spin states along the x-axis (see (12.3.14)): $|x;\pm\rangle = \frac{1}{\sqrt{2}}\big(|+\rangle \pm |-\rangle\big)$. Acting on the plus state with T,

$$T\,|x;+\rangle = \frac{1}{\sqrt{2}}\big(\alpha|-\rangle + \beta|+\rangle\big) = \frac{\beta}{\sqrt{2}}\big(|+\rangle + \tfrac{\alpha}{\beta}|-\rangle\big). \tag{13.3.30}$$

For the result to point along $|x;-\rangle$, we need $\alpha/\beta = -1$. Now consider spin states along the y-axis (see (12.3.14)): $|y;\pm\rangle = \frac{1}{\sqrt{2}}\big(|+\rangle \pm i|-\rangle\big)$. Acting on the plus state with T, we get

$$T\,|y;+\rangle = \frac{1}{\sqrt{2}}\big(\alpha|-\rangle + i\beta|+\rangle\big) = \frac{i\beta}{\sqrt{2}}\big(|+\rangle - i\tfrac{\alpha}{\beta}|-\rangle\big). \tag{13.3.31}$$

For the result to point along $|y;-\rangle$ this time, we need $\alpha/\beta = +1$. The inconsistent constraints on α/β demonstrate that we *cannot* build a linear operator T that reverses the directions of all spin states. There is a basic reason why this operator does not exist. As we will learn later, on a complex vector space any linear operator has at least one eigenvalue and one eigenvector. The eigenvalue cannot be zero since by definition T does not kill spin states. But a nonzero eigenvalue implies an eigenvector, thus a vector that acted by T is just multiplied by the eigenvalue. Such vector a does not change direction, showing no linear operator T can reverse all spin states. There is a map that flips all spin states, but it is not a linear operator.

□