$\square$

14.3 Orthogonal Projectors

We now turn to the definition and construction of *orthogonal projectors*. In general, projectors in some vector space V are linear operators that acting on arbitrary vectors give us a vector in some subspace U of V. In other words, the range of the operator is U. Orthogonal projectors do this but also more: they give zero acting on vectors that are orthogonal to any vector on U. We will discuss this in detail now. A particularly simple characterization of an orthogonal projector will wait for section 14.5, after we discuss the adjoint of an operator. We will show that orthogonal projectors are in fact Hermitian operators (theorem 14.5.4).

We begin our work by noting that an inner product can help us construct interesting subspaces of a vector space V. Consider any *subset* U of vectors in V. Then we can define a *subspace* $U^{\perp}$, called the **orthogonal complement** of U as the set of all vectors orthogonal to the vectors in U:

$$U^{\perp} = \{v \in V \,|\, \langle v, u\rangle = 0, \text{ for all } u \in U\}. \tag{14.3.1}$$

This is clearly a subspace of V. Something remarkable happens when the set U, rather than being just a subset, is itself a *subspace*. In that case, U and $U^\perp$ actually give a direct sum decomposition of the full space:

Theorem 14.3.1. *If U is a subspace of V, then $V = U \oplus U^\perp$.*

Proof. This is a fundamental result and is actually not hard to prove. Let $(e_1, \ldots, e_n)$ be an orthonormal basis for the subspace U. We can then easily write any vector v in V as a sum of a vector in U and a vector in $U^\perp$:

$$v = \underbrace{(\langle e_1, v\rangle e_1 + \cdots + \langle e_n, v\rangle e_n)}_{\in U} + \underbrace{(v - \langle e_1, v\rangle e_1 - \cdots - \langle e_n, v\rangle e_n)}_{\in U_\perp}. \tag{14.3.2}$$

On the right-hand side, the first vector in parenthese is clearly in U as it is written as a linear combination of U basis vectors. The second vector is clearly in $U^\perp$: it is orthogonal to all the basis vectors of U and thus orthogonal to any vector in U. To complete the proof, one must show that there is no vector except the zero vector in the intersection $U \cap U^\perp$ (recall the comments below (13.2.3)). Let $v \in U \cap U^\perp$. Then v is in U and in $U^\perp$, so it should satisfy $\langle v, v\rangle = 0$. But then $v = 0$, completing the proof.

□

Given $V = U \oplus U^\perp$, any vector $v \in V$ can be written *uniquely* as

$$v = u + w, \quad \text{with } u \in U \text{ and } w \in U^\perp. \tag{14.3.3}$$

One can define a linear operator P_U, called the **orthogonal projection** of V onto U, that acting on v above gives the vector u: $P_U v = u$.

It follows from this definition that

$$\begin{aligned} &P_U u = u, \text{ for } u \in U, \\ &P_U w = 0, \text{ for } w \in U^\perp. \end{aligned} \tag{14.3.4}$$

Indeed, for the first we write $u = u + 0$ with $0 \in U^\perp$, and for the second we write $w = 0 + w$ with $0 \in U$. In addition, we have the other following properties:

1. range $P_U = U$.

   The definition of P_U implies that range $P_U \subset U$. The first line in (14.3.4) shows that range $P_U \supset U$. These two together prove the claim. If U is a proper subspace of V (that is, $U \neq V$), P_U is not surjective.

2. null $P_U = U_\perp$.

The second line in (14.3.4) shows that null $P_U \supset U^\perp$. On the other hand, if $\mu \in$ null P_U, we have $P_U\mu = 0$, and by (14.3.3), $\mu = 0 + \mu$, with $0 \in U$, and $\mu \in U^\perp$. This means null $P_U \subset U^\perp$, thus establishing the claim. If U is a proper subspace of V, P_U is not invertible.

3. $P_U^2 = P_U$.

The first line in (14.3.4) tells us that P_U acting on U is the identity operator. Thus, if we act twice with P_U on any vector, the second action has no effect as it is acting on a vector in U. More explicitly, for $v = u + w$ with $u \in U$ and $w \in U^\perp$, we find that

$$P_U P_U v = P_U(P_U v) = P_U\, u = u = P_U v. \tag{14.3.5}$$

Since v is arbitrary, this establishes the claim.

4. With $(e_1, \ldots, e_n)$ an orthonormal basis for U, the action of the projector is explicitly obtained from (14.3.2):

$$P_U v \;=\; \langle e_1, v\rangle e_1 + \cdots + \langle e_n, v\rangle e_n. \tag{14.3.6}$$

5. $\|P_U v\| \le \|v\|$.

The action of P_U cannot increase the length of a vector. Using the decomposition (14.3.3) and the Pythagorean theorem, we find that

$$\|v\|^2 = \|u+w\|^2 = \|u\|^2 + \|w\|^2 \ge \|u\|^2 = \|P_U v\|^2\,. \tag{14.3.7}$$

The claim follows by taking the square root.

Given the results in (1) and (2), the relation $V = U \oplus U^\perp$ becomes

$$V = \text{range}\, P_U \oplus \text{null}\, P_U\,, \tag{14.3.8}$$

and for all $u \in$ range P_U and all $w \in$ null P_U, we have $\langle w, u\rangle = 0$.

Remarks: For any linear operator $T \in \mathcal{L}(V)$, the dimensions of null T and range T add up to the dimension of V (see (13.4.6)). But for an arbitrary T, as opposed to an orthogonal projector, the spaces null T and range T typically have a nonzero intersection and thus do not provide a direct sum decomposition of V. Orthogonal projectors do. The $\hat{S}_\pm$ operators in example 13.14 are operators for which the null space and the range are in fact the same!

The eigenvalues and eigenvectors of P_U are easy to describe. Since all vectors in U are left invariant by the action of P_U, an orthonormal basis for U provides a set of orthonormal eigenvectors of P all with eigenvalue one. If we choose an orthonormal basis for $U^\perp$, that basis provides orthonormal eigenvectors of P all with eigenvalue zero.

In fact, $P_U P_U = P_U$ implies that the eigenvalues of P_U can only be zero or one. Recall that the eigenvalues of an operator satisfy whatever equation the operator satisfies. Therefore, $\lambda^2 = \lambda$ holds, and $\lambda = 0, 1$, are the only possible eigenvalues.

A matrix representation of orthogonal projectors helps us to understand how the projector encodes the dimension of the space U it projects onto. Consider a vector space $V = U \oplus U^\perp$ that is $(n + k)$-dimensional, where U is n-dimensional, and $U^\perp$ is k-dimensional. Let $(e_1, \ldots, e_n)$ be an orthonormal basis for U and $(f_1, \ldots, f_k)$ an orthonormal basis for $U^\perp$. We then see that the list of vectors

$$(e_1, \ldots, e_n, f_1, \ldots, f_k) \text{ is an orthonormal basis for } V. \tag{14.3.9}$$

Since $P_U e_i = e_i$, for $i = 1, \ldots, n$, and $P_U f_j = 0$ for $j = 1, \ldots, k$, these orthonormal basis vectors are in fact eigenvectors of P_U. It follows that in this basis the projector operator is represented by the diagonal matrix:

$$P_U = \operatorname{diag}(\underbrace{1, \ldots, 1}_{n \text{ entries}} \; \underbrace{0, \ldots, 0}_{k \text{ entries}}). \tag{14.3.10}$$

As expected from its noninvertibility, $\det P_U = 0$. More interestingly, the trace of the matrix P_U is n. Therefore,

$$\boxed{\operatorname{tr} P_U = \dim U.} \tag{14.3.11}$$

The dimension of U is the dimension of the range of P_U and thus gives the rank of the projector P_U. Rank-one projectors are the most common projectors. They project to one-dimensional subspaces of the vector space.

Projection operators are useful in quantum mechanics in two ways. First, the act of measuring an observable projects the physical state vector instantaneously to some invariant subspace of the observable; the projection can be seen to be done by an orthogonal projector. Second, orthogonal projectors themselves can be considered observables. These ideas will become clearer once we show that orthogonal projectors are in fact Hermitian operators.

Example 14.5. *Orthogonal projector onto a two-dimensional subspace in* $\mathbb{R}^3$.

Two-dimensional *vector subspaces* of $\mathbb{R}^3$ are planes going through the origin. We can specify planes $\mathcal{C}_{\mathbf{n}}$ using a unit vector $\mathbf{n} = (n_1, n_2, n_3)$ orthogonal to the plane. Then points $\mathbf{x} = (x_1, x_2, x_3)$ on the plane satisfy the constraint

$$\mathbf{n}\cdot\mathbf{x} = n_1x_1 + n_2x_2 + n_3x_3 = 0, \quad \mathbf{x}\in\mathcal{C}_{\mathbf{n}}. \tag{14.3.12}$$

How do we build the orthogonal projector $P_{\mathbf{n}}$ onto $\mathcal{C}_{\mathbf{n}}$? The projector can be thought of as a 3×3 matrix that acting on arbitrary vectors gives the appropriate vector on the plane. At this point it may be a good idea if you try to determine the answer yourself. Then keep reading.

Given any vector $\mathbf{x}$, we rewrite it as follows:

$$\mathbf{x} = \underbrace{\mathbf{x} - (\mathbf{n}\cdot\mathbf{x})\mathbf{n}}_{\mathcal{C}_{\mathbf{n}}} + \underbrace{(\mathbf{n}\cdot\mathbf{x})\mathbf{n}}_{\mathcal{C}_{\mathbf{n}}^{\perp}}. \tag{14.3.13}$$

The first vector on the right-hand side is in $\mathcal{C}_{\mathbf{n}}$ because $\mathbf{n}\cdot(\mathbf{x} - (\mathbf{n}\cdot\mathbf{x})\mathbf{n}) = 0$, recalling that $\mathbf{n}\cdot\mathbf{n} = 1$. The second vector is along $\mathbf{n}$ and thus orthogonal to the plane $\mathcal{C}_{\mathbf{n}}$. The above formula indeed implements the decomposition $\mathbb{R}^3 = \mathcal{C}_{\mathbf{n}} \oplus \mathcal{C}_{\mathbf{n}}^{\perp}$, and the projector $P_{\mathbf{n}}$ acts as

$$P_{\mathbf{n}}\,\mathbf{x} \equiv \mathbf{x} - (\mathbf{n}\cdot\mathbf{x})\mathbf{n}. \tag{14.3.14}$$

We can describe this in matrix notation by looking at the ith component of both sides. With repeated indices summed over,

$$(P_{\mathbf{n}})_{ij}x_j = x_i - (n_jx_j)n_i = (\delta_{ij} - n_in_j)x_j. \tag{14.3.15}$$

From this we read that

$$(P_{\mathbf{n}})_{ij} = \delta_{ij} - n_i n_j = \begin{pmatrix} 1-n_1^2 & -n_1 n_2 & -n_1 n_3 \\ -n_2 n_1 & 1-n_2^2 & -n_2 n_3 \\ -n_3 n_1 & -n_3 n_2 & 1-n_3^2 \end{pmatrix}. \tag{14.3.16}$$

The matrix for $P_{\mathbf{n}}$ is symmetric. Moreover, $P_{\mathbf{n}}$ is a rank-two projector since $\square_{\mathbf{n}}$ is of dimension two. Consistent with this, $\mathrm{tr} P_{\mathbf{n}} = 3 - (n_1^2 + n_2^2 + n_3^2) = 3 - 1 = 2$
.

□

Example 14.6. *Nonorthogonal projector in* $\mathbb{R}^2$.

To better appreciate what an orthogonal projector is, we consider a simple example of a nonorthogonal one. We define a projector P_α taking $\mathbb{R}^2$ to the x_1-axis. Here α is a fixed angle in the range $\alpha \in (0, \frac{\pi}{2})$. For any point (x_1, x_2), draw a line through the point that makes an angle α with the x_1-axis (figure 14.2, *left*). Let x_* denote the coordinate value where the line intersects the x_1-axis. We set

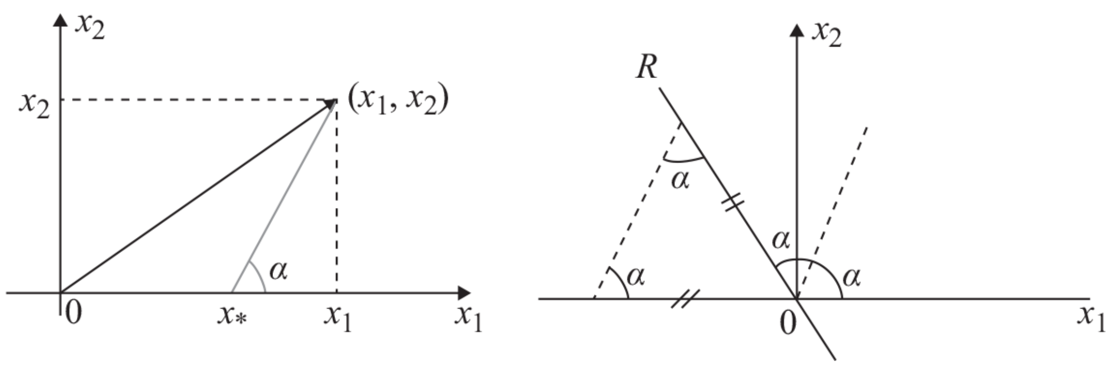

Figure 14.2
Left: The projector P_α takes the point (x_1, x_2) to the point $(x_*, 0)$ on the x_1-axis. *Right*: The projector P_α preserves the length of vectors on the *OR* line, which makes an angle 2α with the x_1-axis. Acting on vectors making an angle greater than 2α and less than π with the x_1-axis, the projector increases their length.

$$P_\alpha \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_* \\ 0 \end{pmatrix}. \tag{14.3.17}$$

A look at the figure shows that

$$x_* = x_1 - \frac{\cos\alpha}{\sin\alpha} x_2. \tag{14.3.18}$$

This implies that

$$P_\alpha = \begin{pmatrix} 1 & -\frac{\cos\alpha}{\sin\alpha} \\ 0 & 0 \end{pmatrix}. \tag{14.3.19}$$

It is simple to see that $P_\alpha P_\alpha = P_\alpha$, as it befits a projector. The null space of P_α is the set of points on the line going through the origin with angle α: $(t\cos\alpha, t\sin\alpha)$, $t \in (-\infty, \infty)$. The trace of the projector matrix is one, equal to the dimensionality of its range. One sign that the projector is not orthogonal is that its matrix is not symmetric—as we already said (but did not show), orthogonal projectors are Hermitian operators. The other sign is that the projector P_α sometimes increases the length of a vector. In figure 14.2, right, the line OR, making an angle of 2α with the x_1-axis, contains the set of points that are projected by P_α with the length unchanged. For points to the "left" of the line OR and above the real line, the projection increases the length.

□

Example 14.7. *Orthogonal projector onto a spin state $|\boldsymbol{n}\rangle$.*

Let $P_{\mathbf{n}}$ project to the one-dimensional space generated by $e_1 = |\mathbf{n}\rangle$. This projector, acting on an arbitrary spin state v, is $P_{\mathbf{n}}v = e_1\langle e_1, v\rangle$. In the notation we have been using for spin states, v is the ket $|v\rangle$ and then $\langle e_1, v\rangle = \langle \mathbf{n}|v\rangle$. As a result, we have

$$P_{\mathbf{n}}|v\rangle = |\mathbf{n}\rangle\langle\mathbf{n}|v\rangle. \tag{14.3.20}$$

Using two-component vectors, we have

$$|\mathbf{n}\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2}e^{i\phi} \end{pmatrix}, \quad \langle\mathbf{n}| = \left(\cos\tfrac{\theta}{2}, \sin\tfrac{\theta}{2}e^{-i\phi}\right), \quad |v\rangle = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}. \tag{14.3.21}$$

The above expression for $P_{\mathbf{n}}$ acting on a state can therefore be rewritten as follows:

$$P_{\mathbf{n}}|v\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2}e^{i\phi} \end{pmatrix} \left(\cos\tfrac{\theta}{2}, \sin\tfrac{\theta}{2}e^{-i\phi}\right) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}. \tag{14.3.22}$$

While conventionally we would multiply the last two factors to form an inner product, it is best to multiply the first two factors, a column vector times a row vector, to find the matrix that represents the action of $P_{\mathbf{n}}$:

$$P_{\mathbf{n}}|v\rangle = \begin{pmatrix} \cos^2\frac{\theta}{2} & \cos\frac{\theta}{2}\sin\frac{\theta}{2}e^{-i\phi} \\ \sin\frac{\theta}{2}\cos\frac{\theta}{2}e^{i\phi} & \sin^2\frac{\theta}{2} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}. \tag{14.3.23}$$

Since the state $|v\rangle$ is arbitrary, the matrix acting on it is the matrix representation of the projector $P_{\mathbf{n}}$. Using double-angle identities, we identify

$$P_{\mathbf{n}} = \frac{1}{2}\begin{pmatrix} 1+\cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & 1-\cos\theta \end{pmatrix} = \frac{1}{2}\mathbb{1} + \frac{1}{2}\begin{pmatrix} n_3 & n_1 - in_2 \\ n_1 + in_2 & -n_3 \end{pmatrix}, \tag{14.3.24}$$

recalling that in spherical coordinates $\mathbf{n} = (n_1, n_2, n_3) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. Finally, we recognize the dot product of $\mathbf{n}$ against the Pauli matrices triplet, so the final result reads

$$P_{\mathbf{n}} = \frac{1}{2}(\mathbb{1} + \mathbf{n}\cdot\boldsymbol{\sigma}). \tag{14.3.25}$$

This is the orthogonal projector we were looking for. It is a rank-one projector.

□