14
Inner Products, Adjoints, and Bra-kets

An inner product is an extra structure that can be added to a vector space. With an inner product, the Schwarz inequality holds, we are able to build orthonormal bases, and we can introduce orthogonal projectors. The inner product allows the definition of a new linear operator: the adjoint $T^\dagger$ *of a linear operator* T*. In an orthonormal basis, the matrix for* $T^\dagger$ *is found by complex conjugation and transposition of the matrix for* T*. An operator is Hermitian if it is equal to its adjoint. Unitary operators are invertible operators that preserve the norm of vectors. A unitary operator* U *satisfies* $UU^\dagger = U^\dagger U = \mathbb{1}$*. We build unitary operators that rotate spin states. We introduce the bra-ket notation of Dirac, where kets represent states and bras are linear functionals on vectors. The notation affords some flexibility and simplifies a number of manipulations.*

14.1 Inner Products

We have been able to go a long way without introducing any additional structure on vector spaces. We have been able to consider linear operators, matrix representations, traces, invariant subspaces, eigenvalues, and eigenvectors. It is now time to put some additional structure on the vector spaces. In this section we consider a function called an *inner product* that allows us to construct numbers from vectors. With inner products we can introduce orthonormal bases for vector spaces and the concept of orthogonal projectors. Inner products will also allow us to define the *adjoint* of an operator. With adjoints available, we can define self-adjoint operators, usually called Hermitian operators in physics. We can also define unitary operators. Some of these ideas were first presented in part I of this book, in less generality and depth.

An **inner product** on a vector space V over a field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$) is a machine that takes an *ordered* pair of elements of V—that is, two vectors—and yields a number in $\mathbb{F}$. A vector space with an inner product is called an *inner-product space*. In order to motivate the definition of an inner product, we first discuss real vector spaces and begin by recalling the way in which we associate a length to a vector in $\mathbb{R}^n$.

The length, or **norm** of a vector, is a real nonnegative number equal to zero if the vector is the zero vector. A vector $a = (a_1, \ldots, a_n)$ in $\mathbb{R}^n$ has norm $\|a\|$ defined by

$$\|a\| = \sqrt{a_1^2 + \cdots + a_n^2}. \tag{14.1.1}$$

Squaring this, we view $\|a\|^2$ as the *dot product* of a with a:

$$\|a\|^2 = a \cdot a = a_1^2 + \cdots + a_n^2. \tag{14.1.2}$$

This suggests that the dot product of any two vectors $a, b \in \mathbb{R}^n$ is defined by

$$a \cdot b \equiv a_1 b_1 + \cdots + a_n b_n. \tag{14.1.3}$$

We now generalize the dot product on $\mathbb{R}^n$ to an inner product, also denoted with a dot, on real vector spaces. This inner product is required to satisfy the following properties:

1. $a \cdot a \geq 0$, for all vectors a.
2. $a \cdot a = 0$ if and only if $a = 0$.
3. $a \cdot (b_1 + b_2) = a \cdot b_1 + a \cdot b_2$.
4. $a \cdot (\alpha\, b) = \alpha\, a \cdot b$, with $\alpha \in \mathbb{R}$ a number.
5. $a \cdot b = b \cdot a$.

Along with these axioms, the length or norm $\|a\|$ of a vector a is the positive or zero number defined by relation

$$\|a\|^2 = a \cdot a. \tag{14.1.4}$$

The third property above is additivity on the second entry. Because of the fifth property, commutativity, additivity also holds for the first entry.

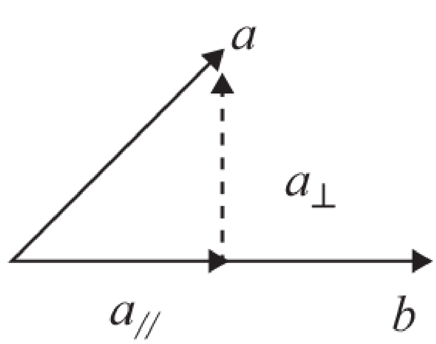
These axioms are satisfied by the definition (14.1.3) but do not require it. A new dot product defined by $a \cdot b = c_1 a_1 b_1 + \cdots + c_n a_n b_n$, with $c_1, \ldots, c_n$ positive constants, would do equally well! (Which axiom goes wrong if we take some c_n equal to zero?) It follows that any result we can prove with these axioms holds true not only for the conventional dot product but for many others as well.

Exercise 14.1. *For the standard, consistent inner product in three dimensions, we have* $a \cdot b = \|a\| \|b\| \cos \theta_{ab}$ *with* θ_{ab} *the angle between the two vectors a and b. Suppose you try to define a new inner product * using half the angle in between the vectors, as in* $a * b = \|a\| \, \|b\| \cos \frac{1}{2}\theta_{ab}$. *Find some reason why this is inconsistent with the axioms (a possible answer will be discussed in example 14.2).*

The above axioms guarantee a fundamental result, the **Schwarz inequality**:

$$|a \cdot b| \leq \|a\| \, \|b\| . \tag{14.1.5}$$

On the left-hand side, the bars denote absolute value. If any of the two vectors is zero, the inequality is trivially satisfied. To prove the inequality, consider two nonzero vectors a and b and then examine the shortest vector joining a point on the line defined by the direction of b to the end of a (figure 14.1). This is the vector $a_\perp$, given by



$$a_\perp = a - a_{//} = a - \frac{(a \cdot b)\, b}{(b \cdot b)}$$

Figure 14.1
The Schwarz inequality follows from the statement that the vector $a_\perp$ must have a nonnegative norm.

$$a_\perp \equiv a - \frac{a\cdot b}{b\cdot b}\, b. \tag{14.1.6}$$

The subscript $\perp$ indicates that the vector is perpendicular to b: $a_\perp \cdot b = 0$, as you can quickly see. To write the above vector, we subtracted from a the component of a parallel to b. Note that the vector $a_\perp$ is not changed as $b \to cb$ with c a constant; it does not depend on the overall length of b. Moreover, as it should, the vector $a_\perp$ is zero if and only if the vectors a and b are parallel.

The Schwarz inequality follows from $a_\perp \cdot a_\perp \geq 0$ as required by axiom (1). Using the explicit expression for $a_\perp$, a short computation gives

$$a_\perp \cdot a_\perp = a\cdot a - \frac{(a\cdot b)^2}{b\cdot b} \geq 0. \tag{14.1.7}$$

Since b is not the zero vector, we then have

$$(a\cdot b)^2 \leq (a\cdot a)(b\cdot b). \tag{14.1.8}$$

Taking the square root of this relation, we obtain the Schwarz inequality (14.1.5). The inequality becomes an equality only if $a_\perp = 0$ or, as discussed above, when $a = cb$ with c a real constant.

For complex vector spaces, some modifications are necessary. Recall that the length $|\gamma|$ of a complex number γ is given by $|\gamma| = \sqrt{\gamma^*\gamma}$, where the asterisk denotes complex conjugation. It is not hard to generalize this a bit. Let $z = (z_1, \ldots, z_n)$ be a vector in $\mathbb{C}^n$. Then the norm $\|z\|$ of z is a real number greater than or equal to zero defined by

$$\|z\|^2 \equiv z_1^* z_1 + \cdots + z_n^* z_n. \tag{14.1.9}$$

We must use complex conjugates to produce a real number greater than or equal to zero. The above suggests that for vectors $w = (w_1, \ldots, w_n)$ and $z = (z_1, \ldots, z_n)$ in $\mathbb{C}^n$, an inner product, denoted by $\langle \cdot, \cdot \rangle$, could be given by

$$\langle w, z\rangle = w_1^* z_1 + \cdots + w_n^* z_n. \tag{14.1.10}$$

Note that we are not treating the two vectors in a symmetric way. There is the first vector, in this case w, whose components are conjugated and a second vector z whose components are not conjugated. If the order of vectors is reversed, the new inner product is the complex conjugate of the original. The order of vectors matters for the inner product in complex vector spaces. We can, however, define an inner product in a way that applies both to complex and real vector spaces. Let us do this now.

An **inner product** on a vector space V over a field $\mathbb{F}$ is a map from an ordered pair (u, v) of vectors in V to a number $\langle u, v\rangle$ in the field. The axioms for $\langle u, v\rangle$ are inspired by the axioms we listed for the dot product:

1. $\langle v, v\rangle \geq 0$, for all vectors $v \in V$.
2. $\langle v, v\rangle = 0$ if and only if $v = 0$.
3. $\langle u, v_1 + v_2\rangle = \langle u, v_1\rangle + \langle u, v_2\rangle$. Additivity in the second entry.
4. $\langle u, \alpha v\rangle = \alpha \langle u, v\rangle$, with $\alpha \in \mathbb{F}$. Homogeneity in the second entry.
5. $\langle u, v\rangle = \langle v, u\rangle^*$. Conjugate exchange symmetry.

The **norm** $\|v\|$ of a vector $v \in V$ is defined by relation

$$\|v\|^2 = \langle v, v\rangle. \tag{14.1.11}$$

Comparing with the dot product axioms for real vector spaces, the key difference is in (5): the inner product in complex vector spaces is not symmetric. For the above axioms to apply to vector spaces over $\mathbb{R}$, we just define the obvious: complex conjugation of a real number is the same real number. On a real vector space, complex conjugation has no effect, and the inner product is strictly symmetric.

Let us make a few remarks. One can use axiom (3) with $v_2 = 0$ to show that for all $u \in V$,

$$\langle u, 0\rangle = 0 \quad \Rightarrow \quad \langle 0, u\rangle = 0, \tag{14.1.12}$$

where the second equation follows by axiom (5). Axioms (3) and (4) amount to full linearity in the second entry. It is important to note that additivity holds for the first entry as well:

$$\begin{aligned}\langle u_1 + u_2, v\rangle &= \langle v, u_1 + u_2\rangle^* \\ &= (\langle v, u_1\rangle + \langle v, u_2\rangle)^* \\ &= \langle v, u_1\rangle^* + \langle v, u_2\rangle^* \\ &= \langle u_1, v\rangle + \langle u_2, v\rangle.\end{aligned} \tag{14.1.13}$$

Homogeneity works differently on the first entry, however:

$$\langle \alpha\, u,\ v\rangle = \langle v, \alpha\, u\rangle^* = (\alpha \langle v, u\rangle)^* = \alpha^* \langle u,\ v\rangle. \tag{14.1.14}$$

In summary, we get linearity and *conjugate homogeneity* on the first entry:

$$\boxed{\begin{aligned} \langle u_1 + u_2, v\rangle &= \langle u_1, v\rangle + \langle u_2, v\rangle, \\ \langle \alpha\, u,\ v\rangle &= \alpha^* \langle u,\ v\rangle. \end{aligned}} \tag{14.1.15}$$

For a real vector space, conjugate homogeneity is just plain homogeneity.

Two vectors $u, v \in V$ are said to be **orthogonal** if $\langle u, v\rangle = 0$. This, of course, means that $\langle v, u\rangle = 0$ as well. The zero vector is orthogonal to all vectors, including itself.

The inner product we have defined is **nondegenerate**: any vector orthogonal to all vectors in the vector space must be equal to zero. Indeed, if $x \in V$ is such that $\langle x, v\rangle = 0$ for all v, pick $v = x$ so that $\langle x, x\rangle = 0$ implies $x = 0$ by axiom (2).

The **Pythagorean** identity holds for the norm squared of orthogonal vectors in an inner-product vector space. As you can quickly verify,

$$\|u + v\|^2 \ = \ \|u\|^2 + \|v\|^2\,, \quad \text{for } u, v \in V, \text{ orthogonal vectors: } \langle u, v\rangle = 0. \tag{14.1.16}$$

The **Schwarz inequality** can be proven by an argument fairly analogous to the one we gave above for real vector spaces. The result now reads

$$\boxed{\text{Schwarz inequality:} \qquad |\langle u, v\rangle| \ \leq \ \|u\| \ \|v\|\,.} \tag{14.1.17}$$

The inequality is saturated if and only if one vector is a multiple of the other. You will prove this identity in a slightly different way in problem 14.1. You will also consider there the **triangle inequality**:

$$\|u + v\| \leq \|u\| + \|v\|\,, \tag{14.1.18}$$

which is saturated when $u = cv$ for c, a real, positive constant. Our definition (14.1.11) of a norm on a vector space V is mathematically sound: a norm is required to satisfy the triangle inequality. Other properties are required: (i) $\|v\| \geq 0$ for all v, (ii) $\|v\| = 0$ if and only if $v = 0$, and (iii) $\|cv\| = |c|\|v\|$ for any constant c. Our norm satisfies all of them.

A finite-dimensional complex vector space with an inner product is a **Hilbert space**. Our study of quantum mechanics will often involve *infinite-dimensional* vector spaces. An infinite-dimensional complex vector space with an inner product is a Hilbert space if an additional *completeness* property holds: all Cauchy sequences of vectors must converge to vectors in the space. An infinite sequence of vectors v_i, with $i = 1, 2, \ldots, \infty$ is a Cauchy sequence if for any $\epsilon > 0$ there is an N such that $\|v_n - v_m\| < \epsilon$ whenever $n, m > N$. For the infinite-dimensional vector spaces we have to deal with, the extra condition holds, and we will not have to concern ourselves with it. We often denote Hilbert spaces with the symbol $\mathcal{H}$.

Example 14.1. *Inner product in* $\mathbb{C}^2$ *for spin one-half.*
We had a first look at spin one-half states and their inner product in sections 12.1 and 12.3. The inner product is described by

$$\langle \Psi, \Phi \rangle = \psi_1^* \phi_1 + \psi_2^* \phi_2, \quad \Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \quad \Phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}, \quad \psi_1, \psi_2, \phi_2, \phi_2 \in \mathbb{C}. \tag{14.1.19}$$

This definition obeys all five axioms of the inner product, as you can quickly verify. In bra-ket notation the inner product is described as

$$\langle \Psi | \Phi \rangle = \psi_1^* \phi_1 + \psi_2^* \phi_2, \quad \text{for} \quad |\Psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \quad |\Phi\rangle = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}. \tag{14.1.20}$$

Recall that we defined $\langle \Psi | \Phi \rangle \equiv \langle \Psi, \Phi \rangle$. While the above inner product is natural and obvious for a complex vector space, the intuition behind it is surprising when using the vectors that represent the direction of spin states, which we consider next.

□

Example 14.2. *Inner product of spin states.*
We have defined the spin state $|\mathbf{n}\rangle \in \mathbb{C}^2$ as the normalized eigenstate of $\hat{\mathbf{S}} \cdot \mathbf{n}$ with eigenvalue $\hbar/2$. We think of this spin state as the state of a particle whose spin vector points in the direction of the unit vector $\mathbf{n}$ (we are using the notation $|\mathbf{n}\rangle = |\mathbf{n}; +\rangle$ and $|-\mathbf{n}\rangle = |\mathbf{n}; -\rangle$). In fact, any spin state in $\mathbb{C}^2$ is, up to a multiplicative constant, a state $|\mathbf{n}\rangle$ for some direction $\mathbf{n}$. This notation associates to any vector in spin space $\mathbb{C}^2$ a vector in ordinary

space $\mathbb{R}^3$. Some confusion can arise because the natural inner products on the two spaces are related in a surprising way. In particular, given two spin states $|\mathbf{n}\rangle$ and $|\mathbf{n}'\rangle$, their inner product in $\mathbb{C}^2$ is *not* given by the inner (dot) product $\mathbf{n} \cdot \mathbf{n}'$ in $\mathbb{R}^3$.

For example, the states $|\mathbf{n}\rangle$ and $|-\mathbf{n}\rangle$ are orthogonal spin states, just like the states $|+\rangle$ and $|-\rangle$. Nevertheless, in $\mathbb{R}^3$ the vectors $\mathbf{n}$ and $-\mathbf{n}$ are antiparallel, and their inner product is nonzero: $\mathbf{n} \cdot (-\mathbf{n}) = -1$. Similarly, spin states along the z-axis and along the x-axis are not orthogonal, while vectors along those directions are.

We can do a simple computation of an inner product between the state $|+\rangle$ whose spin points along the positive z-axis and a state whose vector $\mathbf{n}$ lies on the (x, z) plane and is defined by spherical angles θ and $\phi = 0$. Using the expression for spin states in (12.3.28), we get

$$|\mathbf{n}\rangle = \cos\tfrac{\theta}{2}\,|+\rangle + \sin\tfrac{\theta}{2}|-\rangle. \tag{14.1.21}$$

It follows that

$$|\langle +|\mathbf{n}\rangle|^2 = \cos^2\tfrac{\theta}{2}. \tag{14.1.22}$$

This formula is instructive, as the inner product of the two states is related to the cosine of *half* the angle between the spin vectors in $\mathbb{R}^3$. As you will show in problem 14.2, this result holds for arbitrary spin states:

$$|\langle \mathbf{n}'|\mathbf{n}\rangle|^2 = \frac{1+\mathbf{n}\cdot\mathbf{n}'}{2} = \cos^2\tfrac{\gamma}{2}, \tag{14.1.23}$$

where γ is the angle between the two unit vectors in $\mathbb{R}^3$: $\cos\gamma = \mathbf{n} \cdot \mathbf{n}'$. The above formula demonstrates that the dot product in $\mathbb{R}^3$ can be used to compute the *absolute* value of inner products in spin space, but, of course, it does not determine the possibly complex inner products themselves. We asked in exercise 14.1 if one could define a dot product $*$ in $\mathbb{R}^3$ by using the cosine of half the angle formed by the vectors, as is tempting from the above result. This is not possible: since a and $-a$ are vectors that form an angle of π and $\cos\frac{\pi}{2} = 0$, we would have $a * (-a) = 0$, in contradiction with $a * (-a) = -(a * a) < 0$ for any nonzero vector a.

□

Example 14.3. *Inner product for operators.*

We have considered the vector space $\mathcal{L}(V)$ formed by the linear operators that act on a vector space V. In this vector space, the operators are now the "vectors." Given two operators $A, B \in \mathcal{L}(V)$, we now define an inner product $\langle A, B\rangle \in \mathbb{C}$ satisfying all the desired axioms. In order to do so, we think of the operators as matrices in some arbitrary basis. Inspiration for a definition comes from an attempt to define the norm squared of a single operator A:

$$\langle A, A\rangle = \|A\|^2 . \tag{14.1.24}$$

Imagine a matrix A with all of its entries. We want the norm squared to be such that when it vanishes it sets to zero every single entry—this is what is needed for the matrix to be set to zero, as required by the axioms. There is a simple way to do this. We declare the norm proportional to the sum of squares of the absolute values of each of the entries A_{ij}:

$$\langle A, A\rangle = \frac{1}{2}\sum_{i,j} |A_{ij}|^2 = \frac{1}{2}\sum_{i,j} A_{ij}^* A_{ij}. \tag{14.1.25}$$

This definition certainly does what we wanted. If $\langle A, A\rangle = 0$, each of the entries of the matrix A must vanish. A clearer description of the norm requires a little manipulation in terms of the Hermitian conjugate of the matrix A. Recalling that $A_{ij}{}^* = (A^\dagger)_{ji}$, we have

$$\langle A, A\rangle = \frac{1}{2}\sum_{i,j} (A^\dagger)_{ji} A_{ij} = \frac{1}{2}\sum_{j} (A^\dagger A)_{jj} = \frac{1}{2}\,\mathrm{tr}\,(A^\dagger A). \tag{14.1.26}$$

This now suggests the general definition for the inner product of two operators:

$$\langle A, B\rangle \equiv \frac{1}{2}\,\mathrm{tr}\,(A^\dagger B). \tag{14.1.27}$$

The appearance of the trace is reassuring: we learned that the trace of an operator is basis independent, and therefore, as defined above, the inner product is basis independent. This is as it should be. An example illustrating the above definition will be given in the following section.

□