A short introduction to quantum mechanics IV: inner product in Hilbert space and the orthogonality of eigenvectors of Hermitian matrices

A Hilbert space is a complex vector space where a state vector (or wave function) lives. Here, 'complex' means that coefficients can be complex numbers. In this article, I will show how an inner product (or scalar product) can be defined in a Hilbert space, define Hermitian matrices, and discuss their properties.

First, let's think of a state vector $|v>= (0.6 + 0.8i)|2J>$, where $|2J>$ is an eigenvector of an energy matrix with eigenvalue $2J$. For simplicity, we will assume that the norm of $|2J>$ is 1. Here, 'norm' means the magnitude of a vector; in other words, $\sqrt{<2J|2J>} = 1$ (if you are not familiar with this notation, please read my article "Dirac's bra-ket notation".

Now, let's calculate the norm of the state vector $|v>$ considered above. The norm would be the square root of $<v|v>$. If we naively calculate this, we have $<v| = (0.6 + 0.8i) < 2J|$ and $|v>= (0.6 + 0.8i)|2J>$, so we get $<v|v>= (0.6 + 0.8i)^2 < 2J|2J>= (0.6 + 0.8i)^2$. Therefore the norm we get is $\sqrt{<v|v>} = 0.6 + 0.8i$. This is troublesome as the norm should be a non-negative real number. (There is no such thing as complex number valued magnitude.) However, we can get out of this dilemma by defining the norm of $|v>$ to be the magnitude of $(0.6 + 0.8i)$ times the magnitude of $|2J>$. The magnitude of $(0.6 + 0.8i)$ is $\sqrt{0.6^2 + 0.8^2} = 1$, so the norm of $|v>$ is 1. Another way of writing this is:

$$1 = \sqrt{(0.6 + 0.8i)(0.6 - 0.8i)} = \sqrt{(0.6 + 0.8i)(0.6 + 0.8i)^*} \tag{1}$$

where * denotes complex conjugation. In other words, if we define $<v|$ as $(0.6 + 0.8i)^* < 2J| = (0.6 - 0.8i) < 2J|$, we get

$$<v|v>= (0.6 - 0.8i)(0.6 + 0.8i) < 2J|2J>= 1 \tag{2}$$

Therefore, we conclude that if $|v>= a|w>$, then $<v| = a^* < w|$.

One corollary that follows from this is that $<a|b>^* = <b|a>$. Let's try to prove this rigorously. Let $|a>= \sum_i a_i|i>$ and $|b>= \sum_i b_i|i>$, where $|i>$ is an orthonormal basis. Orthonormality means that $<i|j>= 0$ for i

not equal to j and 1 for i equal to j. Then we get

$$< a|b >= \sum_i (a_i^*) b_i \tag{3}$$

while

$$< b|a >= \sum_i (b_i^*) a_i = \sum_i [(a_i^*) b_i]^*. \tag{4}$$

Therefore we easily see that indeed $< a|b >^* = < b|a >$. One corollary that directly follows from this is that $< v|v >$ is real ($< v|v >^* = < v|v >$).[1]

In a Hilbert space, matrices called Hermitian matrices play a role similar to that played by symmetric matrices in a real vector space. Hermitian matrices are defined by the condition that the complex conjugate of the Hermitian matrix is equal to the transpose of the Hermitian matrix. Or in other words, a matrix that is self-adjoint is called a Hermitian matrix (adjoint is the combination of both transpose and complex conjugation; an operator or a matrix is called self-adjoint if its adjoint is equal to itself). One remarkable property of a Hermitian matrix is that its eigenvalues are always real. Even though we have already learned this from our earlier article "Eigenvalues and eigenvectors of symmetric matrices and Hermitian matrices," we will prove this again using bra-ket notations. Let $|n >$ be an eigenvector of a Hermitian matrix A with eigenvlaue n. Then, we get:

$$< n|A|n >=< n|A^\dagger|n >= (< n|A^\dagger)|n >= (n^* < n|)n >= n^* < n|n > \tag{5}$$

Here, we used the fact that A is self-adjoint. Similarly, we get

$$< n|A|n >=< n|(A|n >) =< n|(n|n >) = n < n|n > \tag{6}$$

Therefore we get

$$n^* < n|n >= n < n|n > \tag{7}$$

This implies that $n^* = n$, which means that n is a real number. Therefore we come to the conclusion that the eigenvalues of a Hermitian matrix are always real. This property of Hermitian matrices is very important. In an earlier article, I mentioned that there is a matrix corresponding to any given observable, and the values for a measurement can only be eigenvalues of this matrix. As the values for the measurement must be real, the eigenvalues must also be real. (It doesn't make any sense that the length of this pencil is "$13 + 3i$" centimeters.) Therefore, we can easily see that the matrices corresponding to observables must be Hermitian matrices.

[1] Actually, in quantum mechanics, we only consider the case in which $\langle v|v\rangle$ is not only real, but also non-negative. However, when you learn quantum field theory or string theory, you will encounter the vectors whose $\langle v|v\rangle$ is not non-negative. Nevertheless, you will learn that such vectors don't correspond to the state vector of actual particles, and are thus called "ghosts." Eliminating ghosts is a big issue in string theory, and this can be only done in certain spacetime dimensions. (26 for bosonic string and 10 for superstring.)

Now, let's prove that all the eigenvectors of a Hermitian matrix are orthogonal to one another. Again, we have proven this before, but we will do this again using bra-ket notation. Let $|n>$ and $|m>$ be eigenvectors of a Hermitain matrix A with eigenvalues n and m. Then, consider $<n|A|m>$. We get:

$$
\begin{aligned}
<n|A|m> &= <n|A^\dagger|m> = (<n|A^\dagger)|m> = (n^* <n|)|m> && (8) \\
&= n* <n|m> = n <n|m> && (9)
\end{aligned}
$$

where we have used the fact n is real, as it is an eigenvalue of a Hermitian matrix ($n^* = n$). Similarly we get

$$<n|A|m> = m <n|m> \tag{10}$$

Equating these results, we get

$$n <n|m> = m <n|m> \tag{11}$$

implying that $<n|m> = 0$ if n is not equal to m. Therefore we have proven that all the eigenvectors of a Hermitian matrix are orthogonal to one another as long as their corresponding eigenvalues are distinct.

This has far-reaching consequences. It implies that we can form an orthogonal basis consisting of eigenvectors of the Hermitian matrices corresponding to observables. (It is known that as long as the eigenvalues are discrete, as opposed to continuous, we can even take the basis to be orthonormal. We will discuss the continuous case later in another article.)

For example letting E_i's be eigenvalues of an Energy matrix, we have an orthonormal basis of $|E_i>$s, that is:

$$
\begin{aligned}
&<E_i|E_j> = 0 \quad \text{if} \quad E_i \neq E_j \\
&<E_i|E_j> = 1 \quad \text{if} \quad E_i = E_j
\end{aligned} \tag{12}
$$

(If originally $<E_i|E_i> = A$ for some non-zero A, then we can 'normalize' by defining a new eigenvector with the same eigenvalue E_i:

$$|E_i(\text{new})> = \frac{|E_i>}{\sqrt{A}} \tag{13}$$

It certainly still has the eigenvalue E_i since

$$E|E_i(\text{new})> = E\left(\frac{|Ei>}{\sqrt{A}}\right) = E_i\frac{|E_i>}{\sqrt{A}} = E_i|E_i(\text{new})> \tag{14}$$

Also, its norm is 1, since

$$<E_i(\text{new})|E_i(\text{new})> = \frac{<E_i|E_i>}{(\sqrt{A})^2} = 1. \tag{15}$$

This is what is meant by 'normalization.'

We can use this orthonormal basis to express an arbitrary vector $|\psi>$ as:

$$|\psi>=\sum_i \psi(E_i)|E_i> \tag{16}$$

This relation can be expressed slightly differently. Recall my article "Diracs bra-ket notation." The completeness relation takes now in the following form:

$$1=\sum_i |E_i><E_i| \tag{17}$$

Multiplying by $|\psi>$ on both-hand side, we conclude:

$$|\psi>=\sum_i <E_i|\psi>|E_i> \tag{18}$$

So, we have $\psi(E_i)=<E_i|\psi>$. In other words, the coordinates of a vector in this basis are just the dot products of the vector with the basis vectors. This is not new. For example, if $v=3\hat{x}+4\hat{y}+5\hat{z}$, we have $5=v\cdot\hat{z}$. This works because our basis is orthonormal.

Problem 1. Let $|a>=(2+3i)|b>$. Obtain $<a|$ in terms of $<b|$.

Problem 2. Assuming $<2\mathrm{J}|2\mathrm{J}>=<3\mathrm{J}|3\mathrm{J}>=1$, check that the following state vector ψ is properly normalized. (i.e. its norm is 1.)

$$|\psi>=\left(\frac{1-i}{\sqrt{6}}\right)|2\mathrm{J}>+\left(\frac{2}{\sqrt{6}}\right)|3\mathrm{J}> \tag{19}$$

Problem 3. If a wave function of a particle is given by the above state vector, what will be the probability of an observer observing its energy to be 2J? What about 3J?

Problem 4. As matrices corresponding to observables must be Hermitian matrices, it is known that the position matrices and the momentum matrices are Hermitian. Given this, check whether the following operators are Hermitian.

$$XP_x, \qquad XY, \qquad XP_y, \qquad P_xP_y \tag{20}$$

Summary

- If $|v\rangle=a|w\rangle$, then $\langle v|=a^*\langle w|$.
- $\langle a|b\rangle^*=\langle b|a\rangle$. Thus, $\langle v|v\rangle$ is always real.
- We can always normalize a non-zero vector to make its norm 1.
- The completeness relation can be expressed as $1=\sum_i |E_i\rangle\langle E_i|$.