October 3, 2001

Recap and linear algebra review

Projection operators

Let's chose an arbitrary pair of kets in a 2-dimensional state space:

$$|\Psi_c\rangle = c_x|x\rangle + c_y|y\rangle \leftrightarrow \begin{pmatrix} c_x \\ c_y \end{pmatrix},$$

$$|\Psi_d\rangle = d_x|x\rangle + d_y|y\rangle \leftrightarrow \begin{pmatrix} d_x \\ d_y \end{pmatrix}. \quad 1$$

Recall that normalization requires $|c_x|^2 + |c_y|^2 = |d_x|^2 + |d_y|^2 = 1$.

The inner product between these two states is given by

$$\langle\Psi_c|\Psi_d\rangle = \begin{pmatrix} c_x^* & c_y^* \end{pmatrix}\begin{pmatrix} d_x \\ d_y \end{pmatrix}$$

$$= c_x^* d_x + c_y^* d_y. \quad 2$$

This inner product is the complex generalization of the familiar "dot product" between vectors in a real Hilbert space:

$$\vec{a} \cdot \vec{b} = a_x b_x + a_y b_y = |a||b|\cos\theta, \quad 3$$

where θ is the included angle.

Since the norms of physical states are required to be 1, we can think of the inner product as defining a kind of generalized "angle" between state vectors:

$$\langle\Psi_c|\Psi_d\rangle \sim \cos\xi. \quad 4$$

In particular, $\langle\Psi_c|\Psi_d\rangle = 0$ implies that the two states are orthogonal, and $\langle\Psi_c|\Psi_d\rangle = 1$ implies that Ψ_c and Ψ_d are the same state. Of course, the inner product between two states is generally a complex number, but we can still think of $|\langle\Psi_c|\Psi_d\rangle|^2$ as being a quantitative measure of the "relative orthogonality" of Ψ_c and Ψ_d.

$\langle\Psi_c|\Psi_d\rangle$ is often called the "overlap" between Ψ_c and Ψ_d.

The projection operator corresponding to a state Ψ_c is given by

$$\mathbf{P}_c = |\Psi_c\rangle\langle\Psi_c|. \quad 5$$

Let's look at the expectation value of $\mathbf{P}_c$ with respect to a state Ψ_d:

$$\begin{aligned}\langle\Psi_d|\mathbf{P}_c|\Psi_d\rangle &= \langle\Psi_d||\Psi_c\rangle\langle\Psi_c||\Psi_d\rangle \\ &= (\langle\Psi_c|\Psi_d\rangle)^*(\langle\Psi_c|\Psi_d\rangle) \\ &= |\langle\Psi_c|\Psi_d\rangle|^2. \quad 6\end{aligned}$$

Hence, $\mathbf{P}_c$ is basically an operator that measures the overlap of other states with Ψ_c.

If we just let $\mathbf{P}_c$ act on a state Ψ_d as an operator,

$$\begin{aligned}\mathbf{P}_c|\Psi_d\rangle &= |\Psi_c\rangle\langle\Psi_c|\Psi_d\rangle \\ &\sim \cos\xi|\Psi_c\rangle.\end{aligned} \quad 7$$

In vector-speak, this is something like the "component of Ψ_d along Ψ_c" or the ***projection*** of Ψ_d on Ψ_c. Hence the name.

At this point it should be clear why $\mathbf{P}_c^2 = \mathbf{P}_c$, right?

Spectral decomposition of a normal operator

"Normal" operators are defined as those operators which satisfy

$$[\mathbf{N},\mathbf{N}^\dagger] \equiv \mathbf{N}\mathbf{N}^\dagger - \mathbf{N}^\dagger\mathbf{N} = 0. \quad 8$$

Clearly, Hermitian ($\mathbf{N} = \mathbf{N}^\dagger$) and unitary ($\mathbf{N}^\dagger\mathbf{N} = \mathbf{N}\mathbf{N}^\dagger = \mathbf{1}$) operators are normal.

In a finite-dimensional vector space, every normal operator has a complete set of orthonormal eigenvectors (see Merzbacher §10.1).

If λ_i^q is an eigenvalue of the operator $\mathbf{O}_q$ and $|i^q\rangle$ is an associated eigenstate,

$$\mathbf{O}_q|i^q\rangle = \lambda_i^q|i^q\rangle. \quad 9$$

If $\mathbf{O}_q$ is a normal operator, its eigenstates $\{|i^q\rangle\}$ can serve as a complete set of orthonormal basis kets for the Hilbert space. [n.b., If $\mathbf{O}_q$ has repeated eigenvalues, we have to diagonalize inside the degenerate subspaces in order to get a good set.] That is, ***every*** state in the state space can be decomposed in the form:

$$\begin{aligned}|\Psi_c\rangle &= \sum_i c_i|i^q\rangle, \\ \sum_i |c_i|^2 &= 1.\end{aligned} \quad 10$$

Let's think about the action of $\mathbf{O}_q$ in this basis:

$$\begin{aligned}\mathbf{O}_q|\Psi_c\rangle &= \mathbf{O}_q\sum_i c_i|i^q\rangle \\ &= \sum_i c_i\mathbf{O}_q|i^q\rangle \\ &= \sum_i c_i\lambda_i^q|i^q\rangle.\end{aligned} \quad 11$$

Likewise, let's think of the action of the projection operators

$$\mathbf{P}_i^q \equiv |i^q\rangle\langle i^q| \quad 12$$

in this basis:

$$
\begin{aligned}
\mathbf{P}_i^q|\Psi_c\rangle &= |i^q\rangle\langle i^q|\sum_{i'} c_{i'}|i'^q\rangle \\
&= \sum_{i'} c_{i'}|i^q\rangle\langle i^q\|i'^q\rangle \\
&= \sum_{i'} c_{i'}|i^q\rangle(\langle i^q|i'^q\rangle) \\
&= \sum_{i'} c_{i'}|i^q\rangle\delta(i,i') \\
&= c_i|i^q\rangle.
\end{aligned} \tag{13}
$$

Hence, we see that

$$
\mathbf{O}_q|\Psi_c\rangle = \sum_i \lambda_i^q \mathbf{P}_i^q|\Psi_c\rangle. \tag{14}
$$

Since this is true for any state Ψ_c, and since the algebra we have just done is basis-independent, we have that

$$
\mathbf{O}_q = \sum_i \lambda_i^q \mathbf{P}_i^q. \tag{15}
$$

This decomposition is called the ***spectral decomposition*** of the operator $\mathbf{O}_q$.

Back to matrix notation

Let's go through an example in two-dimensional Hilbert space, mainly in order to illustrate how the above algebra works in matrix notation. This example may also convince you that bras and kets do a really good job of hiding messy details regarding changes of basis...

Suppose we choose a pair of basis kets for the Hilbert space, $|x\rangle$ and $|y\rangle$. Hence, an arbitrary state $|\Psi_c\rangle$ can be decomposed into

$$
|\Psi_c\rangle = c_x|x\rangle + c_y|y\rangle, \tag{16}
$$

with $|c_x|^2 + |c_y|^2 = 1$. Once we agree that the basis kets are ***fixed***, we can introduce the vector notation

$$
|\Psi_c\rangle \leftrightarrow \begin{pmatrix} c_x \\ c_y \end{pmatrix}. \tag{17}
$$

Now suppose we have a normal operator $\mathbf{O}$, with eigenvalues λ_a and λ_b, and corresponding (orthonormal) eigenstates

$$
\begin{aligned}
|a\rangle &= a_x|x\rangle + a_y|y\rangle \leftrightarrow \begin{pmatrix} a_x \\ a_y \end{pmatrix}, \\
|b\rangle &= b_x|x\rangle + b_y|y\rangle \leftrightarrow \begin{pmatrix} b_x \\ b_y \end{pmatrix}.
\end{aligned} \tag{18}
$$

We can derive the matrix representation (in the x,y basis) for the projection operators corresponding to these eigenvectors:

$$\mathbf{P}_a = |a\rangle\langle a| = \begin{pmatrix} a_x \\ a_y \end{pmatrix}\begin{pmatrix} a_x^* & a_y^* \end{pmatrix} = \begin{pmatrix} |a_x|^2 & a_x a_y^* \\ a_y a_x^* & |a_y|^2 \end{pmatrix},$$

$$\mathbf{P}_b = |b\rangle\langle b| = \begin{pmatrix} b_x \\ b_y \end{pmatrix}\begin{pmatrix} b_x^* & b_y^* \end{pmatrix} = \begin{pmatrix} |b_x|^2 & b_x b_y^* \\ b_y b_x^* & |b_y|^2 \end{pmatrix}. \qquad 19$$

We can thus compute the matrix representation of **O** via spectral decomposition:

$$\mathbf{O} \leftrightarrow \lambda_a \begin{pmatrix} |a_x|^2 & a_x a_y^* \\ a_y a_x^* & |a_y|^2 \end{pmatrix} + \lambda_b \begin{pmatrix} |b_x|^2 & b_x b_y^* \\ b_y b_x^* & |b_y|^2 \end{pmatrix}. \qquad 20$$

Diagonal form of a normal operator

Another interesting way to compute the matrix representation is by constructing the "diagonal form" of **O**. Using O to denote the matrix representation of the operator **O**, we can write

$$O\begin{pmatrix} a_x \\ a_y \end{pmatrix} = \lambda_a \begin{pmatrix} a_x \\ a_y \end{pmatrix},$$

$$O\begin{pmatrix} b_x \\ b_y \end{pmatrix} = \lambda_b \begin{pmatrix} b_x \\ b_y \end{pmatrix}. \qquad 21$$

Hence,

$$O\begin{pmatrix} a_x & b_x \\ a_y & b_y \end{pmatrix} = \begin{pmatrix} \lambda_a a_x & \lambda_b b_x \\ \lambda_a a_y & \lambda_b b_y \end{pmatrix} = \begin{pmatrix} a_x & b_x \\ a_y & b_y \end{pmatrix}\begin{pmatrix} \lambda_a & 0 \\ 0 & \lambda_b \end{pmatrix}. \qquad 22$$

Note that the matrix whose columns are the orthnormal eigenvectors (let's call it S) of a normal operator is itself *unitary*:

$$\begin{pmatrix} a_x^* & a_y^* \\ b_x^* & b_y^* \end{pmatrix}\begin{pmatrix} a_x & b_x \\ a_y & b_y \end{pmatrix} = \begin{pmatrix} |a_x|^2 + |a_y|^2 & a_x^* b_x + a_y^* b_y \\ a_x b_x^* + a_y b_y^* & |b_x|^2 + |b_y|^2 \end{pmatrix} = \mathbf{1}, \qquad 23$$

$$S^\dagger = S^{-1}.$$

Hence, we can multiply equation (22) by S^{-1} from the right, to yield

$$O = S\Lambda S^{-1}, \qquad 24$$

where

$$\Lambda = \begin{pmatrix} \lambda_a & 0 \\ 0 & \lambda_b \end{pmatrix} \qquad 25$$

is the diagonal matrix of eigenvalues. The existence of the factorization (24) for every normal matrix O, with S a unitary matrix whose columns are the eigenvectors of O, is called the Spectral Theorem. Since S can be viewed as a "transformation matrix" that changes the original x, y basis into the a, b basis, this theorem expresses the fact that a matrix is diagonal in its own eigenbasis (provided that basis exists and is complete and orthnormal).

Note that this decomposition is quite useful. For example, it gives us a closed form for arbitrary powers of normal matrices:

$$\begin{aligned} O^n &= (S\Lambda S^{-1})^n \\ &= (S\Lambda S^{-1})(S\Lambda S^{-1})(S\Lambda S^{-1})\cdots \quad \# \\ &= S\Lambda^n S^{-1} \\ &= S\begin{pmatrix} \lambda_a^n & 0 & 0 \\ 0 & \lambda_b^n & 0 \\ 0 & 0 & \ddots \end{pmatrix} S^{-1}. \end{aligned} \tag{26}$$

Of course, we can do a similar trick with our operator version of the spectral decomposition:

$$\begin{aligned} \mathbf{O}^n &= \left(\sum_i \lambda_i |i\rangle\langle i|\right)^n \\ &= \sum_i \lambda_i^n |i\rangle\langle i|. \end{aligned} \tag{27}$$

If you don't see this right away, write it out as an exercise...

Nonorthogonality and imperfect distinguishability

The following simple example is intended to highlight our first genuine example of a mysterious property of quantum mechanics that follows directly from the rules for representation and prediction.

Let's continue to think about our nice simple two-dimensional Hilbert space with basis kets $|x\rangle$ and $|y\rangle$.

Given an arbitrary pair of states

$$\begin{aligned} |\Psi_c\rangle &= c_x|x\rangle + c_y|y\rangle, \\ |\Psi_d\rangle &= d_x|x\rangle + d_y|y\rangle, \end{aligned} \tag{28}$$

under what conditions is it possible to find a (standard) measurement that can distinguish them with zero probability of error?

Recall that a standard measurement is specified by a complete set of orthogonal projectors. Since for now we are working in just two dimensions, we actually only need to specify a single ket $|\phi\rangle$. Then unambiguously,

$$\begin{aligned} \mathbf{P}_1 &= |\phi\rangle\langle\phi|, \\ \mathbf{P}_2 &= \mathbf{1} - \mathbf{P}_1. \end{aligned} \tag{29}$$

In the current scenario, we are trying to pick $|\phi\rangle$ such that

$$\begin{aligned} \Pr(1|c) &= \langle\Psi_c|\mathbf{P}_1|\Psi_c\rangle = 1, \\ \Pr(1|d) &= \langle\Psi_d|\mathbf{P}_1|\Psi_d\rangle = 0, \\ \Pr(2|c) &= \langle\Psi_c|\mathbf{P}_2|\Psi_c\rangle = 0, \\ \Pr(2|d) &= \langle\Psi_d|\mathbf{P}_2|\Psi_d\rangle = 1. \end{aligned} \tag{30}$$

In order to satisfy the first condition, we clearly need to choose

$$|\phi\rangle = |\Psi_c\rangle. \tag{31}$$

As a consequence

$$\Pr(1|d) = \langle \Psi_d|\mathbf{P}_1|\Psi_d\rangle = |c_x^* d_x + c_y^* d_y|^2, \qquad (32)$$

and we find that our two states Ψ_c and Ψ_d are not perfectly distinguishable ***unless*** they have zero overlap:

$$\Pr(1|d) = 0 \quad \textbf{iff} \quad \langle \Psi_c|\Psi_d\rangle = 0. \qquad (33)$$

Hopefully it should be clear that among all possible pairs of vectors in a Hilbert space, only a ***vanishing*** fraction are orthogonal!

And yet, according to our quantum representation rule, every vector in the Hilbert space corresponds to a distinct physical state of the system – that is, to a distinct *preparation* procedure.

Even though quantum measurement theory allows for non-projective measurements (see third term of this course), it is nonetheless a theorem that

- No measurement can distinguish nonorthogonal states, with zero probability of error, in a single trial.

It is a profound mystery that quantum mechanics presents us with such a huge space of possible physical states without allowing us perfectly to distinguish between them. Why is it that we can “put more information into” the preparation of a quantum system than we can “get back out” in a single measurement of that very same system?

In a sense it’s embarassing that we don’t yet have a good answer, but some of the most insightful scientists I know believe that this is the single most important question for contemporary research in quantum theory.