Expectation values and Uncertainty

If we make a large number of measurements of the same observable on equivalent states, we will obtain a distribution of possible results governed by these probabilities. The average of these results, known as the *expectation value* of the observable is denoted by $\bar{\mathcal{O}}$ or $\langle \mathcal{O} \rangle$ and equal to

$$\bar{\mathcal{O}} = \sum_n p_n \lambda_n = \sum_n \lambda_n |\langle \lambda_n | \Psi \rangle|^2 \tag{1}$$

We didn't talk about this yet in class, but I'm including it here since it fits with expectation values:

We may also be interested in how spread out the distribution of results is. A useful measure of this is to compute the standard deviation of the distribution, defined by taking the average value of $(\mathcal{O} - \bar{\mathcal{O}})^2$ and then taking the square root. In quantum mechanics, we call this the *uncertainty* $\Delta\mathcal{O}$ of the observable $\mathcal{O}$ in the state $|\Psi\rangle$. It is given explicitly by

$$\begin{aligned} \Delta\mathcal{O} &= (\sum_n p_n (\lambda_n - \bar{\mathcal{O}})^2)^{\frac{1}{2}} \\ &= (\sum_n p_n \lambda_n^2 - (\sum_n p_n \lambda_n)^2)^{\frac{1}{2}} \ . \end{aligned} \tag{2}$$

0.1 Operators

Given a Hilbert space, a *linear operator* or simply *operator* is defined as a linear map from the vector space to itself, i.e. a map

$$|v\rangle \to \hat{\mathcal{O}}|v\rangle \tag{3}$$

satisfying

$$\hat{\mathcal{O}}(z_1|v_1\rangle + z_2|v_2\rangle) = z_1\hat{\mathcal{O}}|v_1\rangle + z_2\hat{\mathcal{O}}|v_2\rangle \ . \tag{4}$$

The set of operators has the mathematical structure of an *algebra*. This means that they form a complex vector space (i.e. we can add operators and multiply them by a constant), with the additional structure of being able to multiply operators. This multiplication is defined by

$$(\hat{\mathcal{O}}_1\hat{\mathcal{O}}_2)|v\rangle = \hat{\mathcal{O}}_1(\hat{\mathcal{O}}_2|v\rangle) \ . \tag{5}$$

The algebra of operators includes an "identity" element, the operator **1** that takes every vector to itself.

Matrix representation of an operator

By the linearity property, the action of an operator is completely specified by its action on any basis of vectors: if $|v\rangle = \sum c_n |e_n\rangle$ then $\hat{\mathcal{O}}|v\rangle = \sum c_n \hat{\mathcal{O}}|e_n\rangle$.

Acting on any basis vector, the operator must give some linear combination of basis vectors. We can express this as

$$\hat{\mathcal{O}}|e_n\rangle \equiv \sum_m \mathcal{O}_{mn}|e_m\rangle \; , \tag{6}$$

where

$$\mathcal{O}_{mn} \equiv \langle e_m|\mathcal{O}|e_n\rangle \; . \tag{7}$$

are called the matrix elements of the operator $\hat{\mathcal{O}}$.

The matrix $\mathcal{O}_{mn}$ gives a representation of the operator $\hat{\mathcal{O}}$ in a particular basis. If a vector $|v\rangle$ is represented in the same basis by coefficients c_n, the vector $\mathcal{O}|v\rangle$ is represented by coefficients c'_m given by[1]

$$c'_m = \sum_n \mathcal{O}_{mn} c_n \; ; \tag{8}$$

That is, we write the coefficients c_n as a column vector and multiply by the matrix O_{mn} to get the new coefficients.

It often helps to think of these linear operators as performing some type of geometrical transformation on the vector space. We can have various qualitatively different types including operators that stretch things in various ways, operators that preserve lengths similar to rotations, or operators that project onto some lower dimensional subspace.

Eigenvectors and eigenvalues

The action of an operator may be particularly simple on certain vectors. We say that $|v\rangle$ is an *eigenvector* of $\hat{\mathcal{O}}$ with *eigenvalue* λ if

$$\hat{\mathcal{O}}|v\rangle = \lambda|v\rangle \; . \tag{9}$$

Sometimes there can be different eigenvectors with the same eigenvalue. In this case, any linear combination of these eigenvectors is an eigenvector with this eigenvalue.

For some operators, it is possible to choose a basis of vectors which are all eigenvectors of $\hat{\mathcal{O}}$. In this case, $\hat{\mathcal{O}}$ is represented simply as a diagonal matrix.

[1] To see this, we note that $c'_m = \langle e_m|\mathcal{O}|v\rangle = \langle e_m|\mathcal{O}(\sum_n c_n|e_n\rangle) = \sum_n \mathcal{O}_{mn}c_n$.

Hermitian operators

A particular example of this type with special physical relevance in quantum mechanics is the Hermitian operator. For these operators, it is possible to choose the basis of eigenvectors to be orthonormal, and all of the eigenvalues are real. We can visualize the action of a Hermitian operator as stretching the space along each eigenvector direction by a multiplicative factor given by the eigenvalue.

For a Hermitian operator, if we denote the eigenvalues by λ_n and the eigenvectors by $|\lambda_n\rangle$, the matrix elements in this basis are simply

$$\mathcal{O}_{mn} = \lambda_m \delta_{mn} \ , \tag{10}$$

i.e. a diagonal matrix with the eigenvalues along the diagonal.

In another basis, a Hermitian operator will generally not be diagonal. But we can recognize it as Hermitian since in an orthonormal basis, the *adjoint* of the corresponding matrix (defined as the complex conjugate of the transpose) must be equal to the matrix itself:

$$\mathcal{O}_{nm} = \mathcal{O}^*_{mn} \ . \tag{11}$$

We can express this property in a basis independent language: for an operator $\hat{\mathcal{O}}$, we define the adjoint operator $\hat{\mathcal{O}}^\dagger$ to be the operator satisfying

$$\langle \chi | \hat{\mathcal{O}}^\dagger | \psi \rangle = \langle \psi | \hat{\mathcal{O}} | \chi \rangle^* \ . \tag{12}$$

With this definition, a Hermitian operator is an operator satisfying $\hat{\mathcal{O}}^\dagger = \hat{\mathcal{O}}$. A basic theorem in linear algebra states that this is true if and only if there exists an orthonormal basis of eigenvectors with real eigenvalues.

Properties of the Adjoint

From the definition of the adjoint, we can immediately see the following properties that will be useful below

- $(z\hat{\mathcal{O}})^\dagger = z^* \hat{\mathcal{O}}^\dagger$
- If $|w\rangle = \hat{\mathcal{O}}|v\rangle$, then $\langle w|u\rangle = \langle v|\hat{\mathcal{O}}^\dagger|u\rangle$
- $(\hat{\mathcal{O}}_1 \hat{\mathcal{O}}_2)^\dagger = \hat{\mathcal{O}}_2^\dagger \hat{\mathcal{O}}_1^\dagger$
- If $\hat{\mathcal{O}}|v\rangle = \lambda|v\rangle$, then $\langle v|\mathcal{O}^\dagger|w\rangle = \lambda^* \langle v|w\rangle$

From the third property here, we see that the product of two Hermitian operators is generally not Hermitian. However, it follows from the properties of the adjoint that the combination $i[\hat{\mathcal{A}}, \hat{\mathcal{B}}]$ is Hermitian if $\hat{\mathcal{A}}$ and $\hat{\mathcal{B}}$ are Hermitian.