

Lecture 5, January 27, 2006**Solved Homework** (*Homework for grading is also due today*)

We are told that the probability of a random variable taking on a value between $-x$ and x is

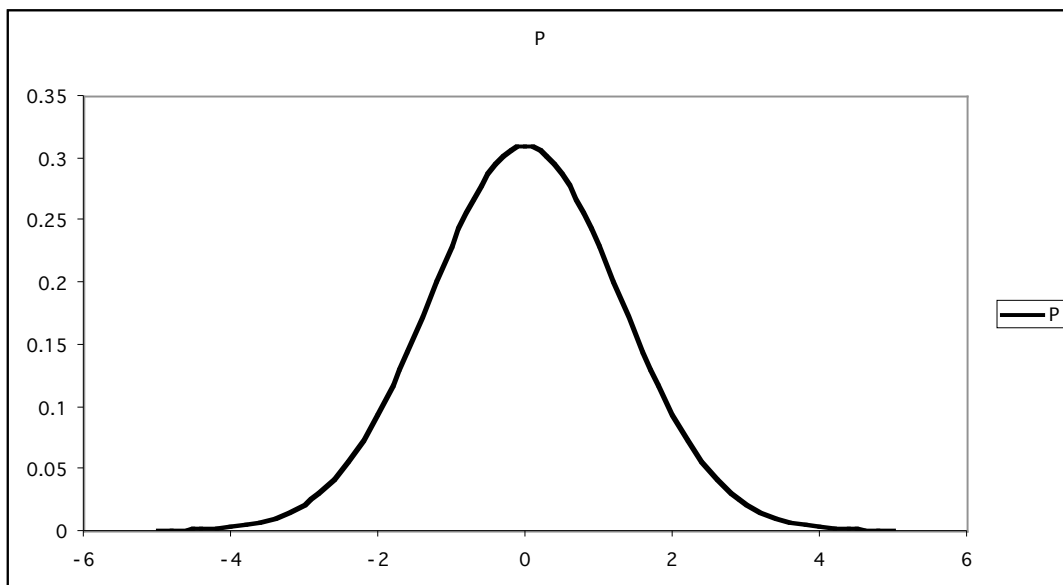
$$P(x) = Ne^{-ax^2}$$

Such a function is called a "gaussian" function, incidentally, and we will see this function many, many times during this course. Since we want P to be normalized, we must have

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} Ne^{-ax^2} dx \\ &= N \int_{-\infty}^{\infty} e^{-ax^2} dx \\ &= N \sqrt{\frac{\pi}{a}} \end{aligned}$$

in which case we demonstrate that N must be $\sqrt{a/\pi}$ to normalize the function (the solution to the integral comes from looking it up in a table).

A sketch for $a = 0.3$, just as an example.



The points of inflection occur where the second derivative is equal to zero.

$$\begin{aligned}\frac{\partial^2 P(x)}{\partial x^2} &= \frac{\partial}{\partial x} \left[\frac{\partial (Ne^{-ax^2})}{\partial x} \right] \\ &= \frac{\partial}{\partial x} (-2axNe^{-ax^2}) \\ &= (4a^2x^2Ne^{-ax^2} - 2aNe^{-ax^2})\end{aligned}$$

setting the r.h.s. equal to zero requires

$$0 = (2ax^2 - 1)$$

so that

$$x = \pm \frac{1}{\sqrt{2a}}$$

For my graph above, where $a = 0.3$, that is $x = \pm 1.29$. (Note that these points *define* one standard deviation for a random variable having “normal” distribution.)

Now, to compute the mean value of x , we must solve

$$\langle x \rangle = \sqrt{\frac{a}{\pi}} \int_{-\infty}^{\infty} xe^{-ax^2} dx$$

while one could go through the math to determine the value of this integral, it is simpler to employ a bit of common sense. From the shape of the probability function, the probability of getting any value x is exactly equal to the probability of getting $-x$. In such a case, the average value of x will necessarily be zero. (Later in the course, we will see that parity arguments can also be used to come to this conclusion without explicitly evaluating the integral.)

The mean value of x^2 , on the other hand, is computed from

$$\langle x^2 \rangle = \sqrt{\frac{a}{\pi}} \int_{-\infty}^{\infty} x^2 e^{-ax^2} dx$$

Integral tables provide

$$\int_{-\infty}^{\infty} x^{2n} e^{-ax^2} dx = \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1)}{2^n a^n} \left(\frac{\pi}{a}\right)^{1/2}$$

In our case, $n = 1$, the square roots cancel one another, and we are left with simply $\langle x^2 \rangle = (2a)^{-1}$. In probability, $\langle x^2 \rangle$ is called σ^2 and we indeed see that the square root is one standard deviation, consistent with our work thus far.

Finally, with respect to the question why is $\langle x^2 \rangle$ not equal to $\langle x \rangle^2$, that is fairly obvious. The former expectation value measures distance from the mean without respect to direction (which is to say, deviation to either left or right is weighted positive by virtue of squaring), while the latter is simply the square of the mean, not the deviation from it. In statistics, the difference between the two is referred to as the “dispersion” of the data.

Necessary Mathematical Tools

Complex numbers. Wave functions and operators can take on complex values, even though *expectation values*, i.e., the results of physical measurements, are always real numbers. Every complex number c can be represented as

$$c = a + bi \tag{5-1}$$

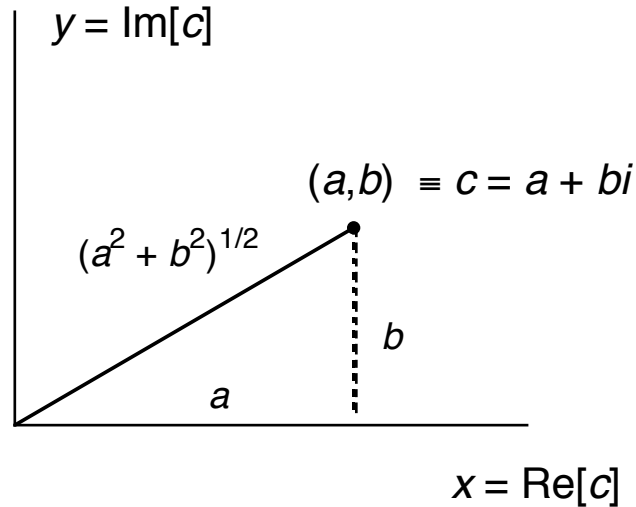
where a and b are real numbers and i is the square root of -1 (the “unit” for the so-called imaginary part of complex numbers). We refer to the “complex conjugate” of c , which is written c^* , as

$$c^* = a - bi \tag{5-2}$$

The “square modulus” of the complex number c is

$$c^* c = |c|^2 = (a + bi)(a - bi) = a^2 + b^2 \tag{5-3}$$

Note that one way to view complex numbers is as vectors in a 2-dimensional cartesian coordinate system. The x axis represents the real component of c (written $\text{Re}[c]$) and the y axis the imaginary component of c (written $\text{Im}[c]$). In that case, it becomes clear that the square modulus of c is simply the square of the length of the corresponding vector in the complex plane. This is illustrated below.



Vectors of complex numbers. Imagine that we have a vector whose elements are complex numbers. Dirac notation in quantum mechanics provides a shorthand connection between vectors and matrix algebra. We refer to a column vector of complex numbers as a “ket”, which is written as $|\tau\rangle$ and defined as

$$|\tau\rangle \equiv (c_1, c_2, c_3, \dots) \equiv \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{bmatrix} \quad (5-4)$$

The term in parentheses is written like a vector specification in an n -dimensional coordinate system (n may be infinite) and the term in brackets is a matrix of n rows and 1 column (also called a column vector).

We may define the analog of the square modulus for such a column vector. In general, the square of the length of a vector may be computed as the product of the transpose of the vector with itself. The transpose of a column vector is defined as the row vector having the same elements (in order) as the column vector. We might write

$$|\tilde{\tau}\rangle \equiv [c_1 \quad c_2 \quad c_3 \quad \dots] \quad (5-5)$$

The transpose definition is true in reverse as well (i.e., the column vector is the transpose of the row vector; in general, the transpose of a matrix with an arbitrary number of rows and columns is defined to be the matrix formed by converting every i, j 'th element of the original matrix to a j, i 'th element of a new matrix).

The rules of matrix multiplication define that when an $m \times n$ matrix (where m defines the number of rows and n the number of columns) multiplies an $n \times l$ matrix (n *must* be the same for both matrices!), then the result is an $m \times l$ matrix. So, for a row vector (1 row, n columns) multiplying its transpose column vector (n rows, 1 column),

one obtains a 1×1 matrix, which is to say a single number, and this defines the square modulus of the vector. Note that as discussed so far, we would have

$$|\bar{\tau}\rangle\langle\tau| = c_1^2 + c_2^2 + c_3^2 + \dots \quad (5-6)$$

This is fine if all the values c are real numbers, but if they are complex numbers, the square of a complex number is still generally complex. To get a real number, we need to use square moduli instead of simple squares. So, for arbitrary kets, we define the “bra”, written $\langle\tau|$, as the complex-conjugate transpose, or *adjoint* of the ket, i.e.,

$$\langle\tau| \equiv [c_1^* \quad c_2^* \quad c_3^* \quad \dots] \quad (5-7)$$

in which case the square modulus of tau is well defined as

$$\langle\tau|\tau\rangle = c_1^*c_1 + c_2^*c_2 + c_3^*c_3 + \dots \quad (5-8)$$

When $\langle\tau|\tau\rangle = 1$, the vector τ is said to be “normalized”. Note that the general notation for the adjoint of an arbitrary matrix is \mathbf{A}^\dagger .

In general, we use this notation because it is convenient for wave functions. Recall that a generic solution to the time-dependent Schrödinger equation is

$$\Psi(x, y, z, t) = \sum_{n=1}^{\infty} c_n \psi_n(x, y, z) e^{-iE_n t / \hbar} \quad (5-9)$$

It is much quicker to write this as $|\Psi\rangle$, where it is implicit in the notation that each i 'th element of the vector is the i 'th element of the sum. To indicate why this is particularly useful, we need to digress for a bit to discuss the properties of quantum mechanical operators.

Hermitian Operators. Recall that an operator A acts on a wave function to deliver the value of an observable a according to

$$A\Psi = a\Psi \quad (5-10)$$

where Ψ is called an eigenfunction of A and a is its associated eigenvalue. Let us now multiply both sides of eq. 5-10 from the left by Ψ^* and integrate to obtain

$$\begin{aligned} \int_{-\infty}^{\infty} \Psi^* (A\Psi) dr &= \int_{-\infty}^{\infty} \Psi^* (a\Psi) dr \\ &= a \int_{-\infty}^{\infty} \Psi^* \Psi dr \\ &= a \end{aligned} \quad (5-11)$$

where dr is the volume element of a space of arbitrary dimensions. Since a is just a number (indeed, a *real* number, since expectation values of observables must be real numbers) it comes outside the integral and normalization of the wave function provides the final result. Indeed, this operationally is *how* one computes a (see last lecture's section on expectation values).

Now, let us consider the special case where it is also true that

$$A^* \Psi = a \Psi \quad (5-12)$$

An operator satisfying both eqs. 5-10 and 5-12 is said to be *Hermitian*. It is trivial to show, then, that the following holds true for a Hermitian operator

$$\int_{-\infty}^{\infty} \Psi^* (A\Psi) dr = \int_{-\infty}^{\infty} \Psi^* (A^* \Psi) dr = \left[\int_{-\infty}^{\infty} \Psi^* (A\Psi) dr \right]^* = \left[\int_{-\infty}^{\infty} \Psi^* (A^* \Psi) dr \right]^* \quad (5-13)$$

Let us consider the case where Ψ is the sum of two arbitrary functions

$$\Psi = \Psi_1 + \Psi_2 \quad (5-14)$$

If we plug this expression into the central equality in eq. 5-13, we have

$$\int_{-\infty}^{\infty} (\Psi_1 + \Psi_2)^* [A^* (\Psi_1 + \Psi_2)] dr = \left[\int_{-\infty}^{\infty} (\Psi_1 + \Psi_2)^* [A(\Psi_1 + \Psi_2)] dr \right]^* \quad (5-15)$$

which may be expanded to

$$\begin{aligned} & \int_{-\infty}^{\infty} \Psi_1^* A^* \Psi_1 dr + \int_{-\infty}^{\infty} \Psi_1^* A^* \Psi_2 dr \\ & + \int_{-\infty}^{\infty} \Psi_2^* A^* \Psi_1 dr + \int_{-\infty}^{\infty} \Psi_2^* A^* \Psi_2 dr \\ & = \left(\int_{-\infty}^{\infty} \Psi_1^* A \Psi_1 dr \right)^* + \left(\int_{-\infty}^{\infty} \Psi_1^* A \Psi_2 dr \right)^* \\ & \quad + \left(\int_{-\infty}^{\infty} \Psi_2^* A \Psi_1 dr \right)^* + \left(\int_{-\infty}^{\infty} \Psi_2^* A \Psi_2 dr \right)^* \end{aligned} \quad (5-16)$$

From the central equality of eq. 5-13, we may eliminate the first and last terms on both sides as they are equal to one another, leaving us with

$$\begin{aligned} & \int_{-\infty}^{\infty} \Psi_1^* A^* \Psi_2 dr + \int_{-\infty}^{\infty} \Psi_2^* A^* \Psi_1 dr \\ & = \left(\int_{-\infty}^{\infty} \Psi_1^* A \Psi_2 dr \right)^* + \left(\int_{-\infty}^{\infty} \Psi_2^* A \Psi_1 dr \right)^* \end{aligned} \quad (5-17)$$

We can now do exactly the same thing, but with a different definition of Ψ , namely

$$\Psi = \Psi_1 + i\Psi_2 \quad (5-18)$$

In this case, we will have

$$\begin{aligned} \Psi^* &= (\Psi_1 + i\Psi_2)^* \\ &= \Psi_1^* - i\Psi_2^* \end{aligned} \quad (5-19)$$

If we follow the same approach used in eq. 5-16, we have

$$\begin{aligned} &\int_{-\infty}^{\infty} \Psi_1^* A^* \Psi_1 dr + i \int_{-\infty}^{\infty} \Psi_1^* A^* \Psi_2 dr \\ &- i \int_{-\infty}^{\infty} \Psi_2^* A^* \Psi_1 dr + \int_{-\infty}^{\infty} \Psi_2^* A^* \Psi_2 dr \\ &= \left(\int_{-\infty}^{\infty} \Psi_1^* A \Psi_1 dr \right)^* + \left[\int_{-\infty}^{\infty} \Psi_1^* A (i\Psi_2) dr \right]^* \\ &\quad - \left(\int_{-\infty}^{\infty} i\Psi_2^* A \Psi_1 dr \right)^* + \left(\int_{-\infty}^{\infty} \Psi_2^* A \Psi_2 dr \right)^* \\ &= \left(\int_{-\infty}^{\infty} \Psi_1^* A \Psi_1 dr \right)^* - i \left[\int_{-\infty}^{\infty} \Psi_1^* A \Psi_2 dr \right]^* \\ &\quad + i \left(\int_{-\infty}^{\infty} \Psi_2^* A \Psi_1 dr \right)^* + \left(\int_{-\infty}^{\infty} \Psi_2^* A \Psi_2 dr \right)^* \end{aligned} \quad (5-20)$$

where it is quite critical to be careful in keeping track of complex conjugate sign changes! Again, we can eliminate the first and last terms on each side based on the central equality of eq. 5-13. If we also divide all remaining terms by i we are left with

$$\begin{aligned} &\int_{-\infty}^{\infty} \Psi_1^* A^* \Psi_2 dr - \int_{-\infty}^{\infty} \Psi_2^* A^* \Psi_1 dr \\ &= - \left[\int_{-\infty}^{\infty} \Psi_1^* A \Psi_2 dr \right]^* + \left(\int_{-\infty}^{\infty} \Psi_2^* A \Psi_1 dr \right)^* \end{aligned} \quad (5-21)$$

Now, if we *add* eq. 5-17 to eq. 5-21, we lose two terms by cancellation and, after eliminating a common factor of 2, we are left with

$$\begin{aligned} \int_{-\infty}^{\infty} \Psi_1^* A^* \Psi_2 dr &= \left(\int_{-\infty}^{\infty} \Psi_2^* A \Psi_1 dr \right)^* \\ &= \int_{-\infty}^{\infty} A^* \Psi_1^* \Psi_2 dr \end{aligned} \quad (5-22)$$

Finally, we know from A being Hermitian that $A^* \Psi_2 = A \Psi_2$, so we may substitute the l.h.s. to arrive at the “turnover rule”, namely

$$\begin{aligned}\int_{-\infty}^{\infty} \Psi_1^* A \Psi_2 dr &= \int_{-\infty}^{\infty} A^* \Psi_1^* \Psi_2 dr \\ &= \int_{-\infty}^{\infty} (A \Psi_1)^* \Psi_2 dr\end{aligned}\quad (5-23)$$

The turnover rule for Hermitian operators is tremendously useful in various manipulations of integral quantities in quantum mechanics.

Eigenvalues and Eigenvectors of Hermitian Operators

Let us apply the turnover rule to a single eigenfunction Ψ of A . Note that

$$\begin{aligned}\int_{-\infty}^{\infty} \Psi^* A \Psi dr &= \int_{-\infty}^{\infty} (A \Psi)^* \Psi dr \\ \int_{-\infty}^{\infty} \Psi^* a \Psi dr &= \int_{-\infty}^{\infty} (a \Psi)^* \Psi dr \\ a \int_{-\infty}^{\infty} \Psi^* \Psi dr &= a^* \int_{-\infty}^{\infty} \Psi^* \Psi dr \\ a &= a^*\end{aligned}\quad (5-23)$$

For a to be equal to a^* , a must be real. This is no surprise, of course, as we *defined* Hermitian operators to have real numbers as eigenvalues, but it is important to emphasize again that this is exactly what we want for quantum mechanical operators, that they produce *only real expectation values*. Hence, we have a fundamental postulate that *all quantum mechanical observables have associated with them a Hermitian operator*.

Let us consider another property of Hermitian operators. If our operator A has two different eigenfunctions with associated eigenvalues according to

$$A \Psi_i = a_i \Psi_i \quad \text{and} \quad A \Psi_j = a_j \Psi_j \quad (5-24)$$

Now, consider multiplication on the left side of the first equality by Ψ_j^* followed by integration, that is

$$\int_{-\infty}^{\infty} \Psi_j^* A \Psi_i dr = a_i \int_{-\infty}^{\infty} \Psi_j^* \Psi_i dr \quad (5-25)$$

And, by the turnover rule

$$\begin{aligned}\int_{-\infty}^{\infty} \Psi_j^* A \Psi_i dr &= \int_{-\infty}^{\infty} (A \Psi_j)^* \Psi_i dr \\ &= a_j \int_{-\infty}^{\infty} \Psi_j^* \Psi_i dr\end{aligned}\quad (5-26)$$

So, if we subtract eq. 5-25 from eq. 5-26, we have

$$(a_j - a_i) \int_{-\infty}^{\infty} \Psi_j^* \Psi_i dr = 0 \quad (5-27)$$

For this equality to be true, either the two eigenvalues a must be equal to one another (that is, the expectation values are "degenerate"), or, the integral must be zero. When the integral is zero, we say that the two functions are "orthogonal". If they are also normalized (remember that it is trivial to normalize a wave function), we say the functions are "orthonormal" and we write

$$\int_{-\infty}^{\infty} \Psi_j^* \Psi_i dr = \delta_{ij} \quad (5-28)$$

where δ is the "Kronecker delta", which is defined as

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (5-29)$$

Note that if eq. 5-27 is satisfied because the eigenvalues are degenerate, we can still construct wave functions using Ψ_i and Ψ_j that will be eigenfunctions of A and yet be orthogonal (this is the homework for next class). Thus, we can state as a proven principle that *the eigenfunctions of a Hermitian operator are orthonormal*. This property will be very useful for simplifying integrals later on.

We may take advantage of this immediately in a notational sense. Consider the case where a wave function is expressed as a linear combination of orthonormal functions. That is,

$$\Phi = \sum_{i=1}^n c_i \Psi_i \quad (5-30)$$

In that case, we would have

$$\begin{aligned} \int_{-\infty}^{\infty} \Phi^* \Phi dr &= \int_{-\infty}^{\infty} \left(\sum_{i=1}^n c_i \Psi_i \right)^* \left(\sum_{j=1}^n c_j \Psi_j \right) dr \\ &= \sum_{i,j=1}^n c_i^* c_j \int_{-\infty}^{\infty} \Psi_i^* \Psi_j dr \\ &= \sum_{i,j=1}^n c_i^* c_j \delta_{ij} \\ &= |c_1|^2 + |c_2|^2 + \cdots + |c_n|^2 \end{aligned} \quad (5-31)$$

Notice that the final result is exactly the same as for multiplication of a column vector by its adjoint row vector when the elements of the vector are the coefficients of the wave function. This is the key connection between our wave function notation thus far and the matrix notation we discussed at the beginning. Thus, in Dirac notation, we simplify our notation by defining

$$\int_{-\infty}^{\infty} \Phi^* \Phi dr = \langle \Phi | \Phi \rangle \quad (5-32)$$

In addition, when there is an operator involved, we may consider the matter from the context of matrix algebra. Our final integral is just a number, so we still want the result of the operator acting on the column vector to be a column vector times a scalar (the scalar is the eigenvalue, see eq. 5-11). This implies that the operator must be a $n \times n$ matrix, since a $n \times n$ matrix multiplies a $n \times 1$ column vector to give a $n \times 1$ column vector. This idea of a matrix representation of an operator will prove useful much later in the class, but for now is simply a point of general interest. More practically for the moment, let us define the Dirac notation

$$\int_{-\infty}^{\infty} \Phi^* A \Phi dr = \langle \Phi | A | \Phi \rangle \quad (5-33)$$

While we've written eqs. 5-32 and 5-33 with a bra that is the specific adjoint of the ket, that need not generally be the case. For instance, we could write eq. 5-28 as

$$\langle \Psi_j | \Psi_i \rangle = \delta_{ij} \quad (5-34)$$

It is critical to remember always that Dirac notation represents integrals, but *those integrals are just numbers* (zero, one, π^8 , 0.31456..., whatever).

Commutators

Notice that when two operators are applied in sequence, the order of the sequence may matter. For instance, if we consider the operators

$$A \equiv x \quad \text{and} \quad B \equiv \frac{\partial}{\partial x} \quad (5-35)$$

and their operation on the function x^3 , we find

$$\begin{aligned} A[B(x^3)] &= A[3x^2] \\ &= 3x^3 \end{aligned} \quad (5-36)$$

and

$$\begin{aligned} B[A(x^3)] &= A[x^4] \\ &= 4x^3 \end{aligned} \quad (5-37)$$

These results are *not* equal.

We define the “commutator” of two operators as

$$[A, B] = AB - BA \quad (5-38)$$

The commutator is *itself* an operator, of course. Indeed, from eqs. 5-36 and 5-37, we have

$$[A, B]x^3 = -x^3 \quad (5-39)$$

or, $[A, B] = -1$ (a rather simple operator, which changes the sign of its operand—it is trivial to prove that this is true for *any* function, not just x^3). From eq. 5-38, it is obvious that $[B, A] = -[A, B]$. When $[A, B] = 0$, the operators A and B are said to “commute”.

Commuting operators have several important properties—here is a critical executive summary:

If Hermitian operators A and B commute, then

$$\langle \Psi_i | B | \Psi_j \rangle = b_i \delta_{ij} \quad (5-40)$$

where the elements of $\{\Psi\}$ are a *necessarily existing* set of orthonormal eigenfunctions that are *common* to both A and B , and the elements of $\{b\}$ are the associated eigenvalues of B (the full set of eigenvalues is called the “spectrum” of B).

Homework

To be solved in class:

Given eqs. 5-10 and 5-12 where a is a real number, prove the equalities in eq. 5-13.

To be turned in for possible grading Feb. 3:

Consider a pair of degenerate, normalized eigenfunctions, ϕ_1 and ϕ_2 , of a Hermitian operator A with common eigenvalue a . Show that two new functions defined as $u_1 = \phi_1$ and $u_2 = \phi_2 - S\phi_1$ are orthogonal, provided that S is properly chosen (i.e., determine what value of S is required to enforce orthogonality). Show that u_1 and u_2 remain degenerate with common eigenvalue a .