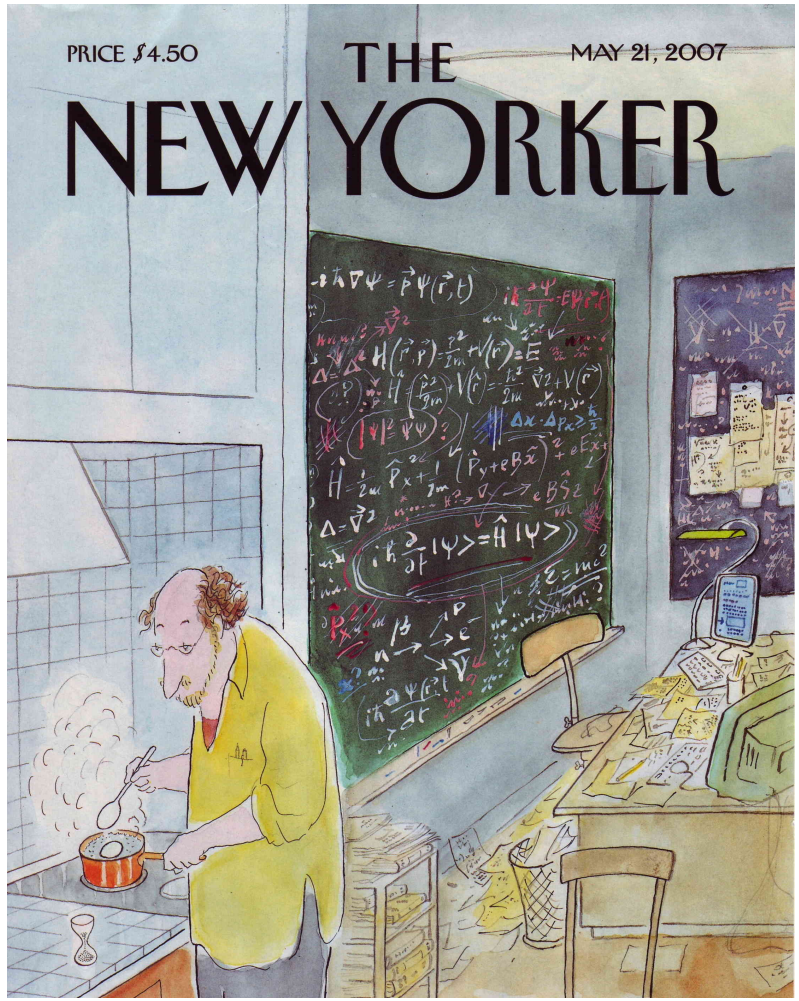
Ph125: Quantum Mechanics

Sunil Golwala

Fall 2008/Winter 2009



Contents

Contents

Contents

Contents

Contents

Lecture 1:
Introduction to Course, Postulates of QM, and Vector Spaces

Date Revised: 2008/09/29
Date Given: 2008/09/29

Section 1

Introduction to Course

Course Logistics

- The course webpage is

  http://www.astro.caltech.edu/~golwala/ph125ab/

  Much of the material will be linked from the course Moodle page, which you can access from the above page or directly via

  https://courses.caltech.edu

  A password for the Ph125a page will be provided in class; you can also obtain it from your classmates, the TAs, or myself. All course logistics and assignments will be announced via the Moodle page. You will find a listing of the course syllabus, problem sets, and solutions there. There is also a weekly homework survey. It would be very beneficial (to me and you) if you could fill out the survey regularly. Especially important is the "News Forum", via which I will make course announcements that I believe you will receive automatically via email once you have logged in to the course page. This is the first time Moodle is in widespread use at Caltech, and the first time I am using it, so please bear with me as I figure it out.

  Comments on the course (or the Moodle page) via the Moodle website are welcome and encouraged. Unfortunately, such comments are not anonymous, so please use campus mail to one of my mailboxes if anonymity is desired.
- Text: Shankar, lecture notes. Many other nice texts are available, choose a different one if you don't like Shankar. See the course reserve list.
- Syllabus: Detailed syllabus on web. Stay on top of it!

Course Logistics (cont.)

- Problem sets: one per week, 4-6 problems. Posted via web site. Due date: **Tuesday, 4 pm**. Problem sets will be much easier to complete in a reasonable amount of time if you stay on top of lectures and clarify any confusion on material *before* you begin the problem set. Solutions posted on web shortly after set is due, graded sets handed back by end of following week. Keep a copy of your problem sets if you want to check your work against the solutions promptly (waiting 1.5 weeks is a bad idea...).
- Grading: 1/3 problem sets (weekly), 1/3 midterm, 1/3 final. (No problem set during week that midterm is due.)
- Exams: each will be 4 hours, 4-6 problems, take home, 1 week lead time. Should only take 2 hours.
- Class attendance: not mandatory. If you don't find it useful, then don't come. The lecture notes will be posted online promptly. But please make the decision based on a careful evaluation of whether you find lecture useful, not on your sleep schedule or time pressure from other classes. And, most importantly, do not think that, just because all the course material is available online, you can catch up the night before the problem set or exam is due and do well. If you don't attend class, be disciplined about doing the reading and following the lecture notes.
- Office hours: I will hold an evening office hour the night before problem sets are due. I strongly prefer Monday night 7-9 pm. There will be a TA office hour also.
- TAs will be primarily responsible for solution sets and grading. They will rotate through the TA office hour.

Course Material

In this course, you will learn to attack basic quantum mechanical problems from scratch and arrive at full solutions that can be tested by experiment. You will see much material that is familiar to you from Ph2/12, but we will cover that material more deeply and with a more formal foundation that will provide you the tools to attack new problems, whether in other courses or in research.

Course Material (cont.)

Prerequisites

Physics:

- Quantum mechanics: None required, in principle. While most students taking this course will have had a course in quantum mechanics before at the level of Ph 2/12, we develop all concepts from scratch and do not require that you recall results from a previous course. However, because we take a formal, systematic approach, basic familiarity with quantum mechanics at the level of Ph 2/12 will be helpful in motivating various parts of the course — essentially, in seeing where we are going. *If you have never had a QM course before at the level of Ph 2/12, you will have to judge for yourself whether you are ready for this course or not.*
- Classical mechanics: Nothing more than Ph1a-level classical mechanics is required. Where we need more sophisticated concepts, we will provide the necessary background material. Knowledge of Hamiltonian mechanics will help in motivating some of the concepts we deal with, but is not necessary and will not be assumed.

Course Material (cont.)

Mathematics:

- Multivariate differential and integral calculus in cartesian and non-cartesian (cylindrical and spherical) coordinate systems at the level of Ph1abc.
- Vectors and vector operations at the level of Ph1abc.
- Methods for solving first- and second-order linear ordinary differential equations at the level of Ph1abc (exponentials and simple harmonic oscillators).
- We will use separation of variables to solve first- and second-order linear partial differential equations, but we will not assume you already know how.
- Linear algebra: We do not assume any prior knowledge of linear algebra aside from matrix multiplication and systems of linear equations (essentially, high-school algebra) along with glancing familiarity with concepts like orthogonal and symmetric matrices. We will develop the necessary more sophisticated concepts here. However, you must quickly become adept in linear algebra because it is the *language* of quantum mechanics. Linear algebra must become second nature to you.
- Key point: Mathematics is the language of physics. You must be competent in above basic mathematical physics in order to understand the material in this course. Intuition is important, but few can succeed in physics without learning to formalize that intuition into mathematical concepts and calculate with it.

Course Material (cont.)

Topics to be covered:

- Mathematical foundations for QM.
- Fundamental postulates of QM: our framework for how we discuss states, physical observables, and interpret quantum states.
- Simple one-dimension problems — building your intuition with piecewise-constant potentials.
- Harmonic Oscillator — the archetypal QM problem.
- Commutations and uncertainty relations — how the noncommutativity of the operators for physical observables results in minimum uncertainties when performing noncommuting measurements.
- Multiparticle systems: Fock product spaces, treatment of systems of identical particles (symmetry/antisymmetry of states).
- Approximate methods for problems without exact solutions: WKB approximation, variational method.
- Classical rotations in three spatial dimensions; tensors.
- Symmetries: esp. how symmetries of the Hamiltonian determine conserved observables.
- Coordinate angular momentum. How to use the angular momentum observables to classify 3D states.

Course Material (cont.)

- Formalism for spin angular momentum.
- Addition of angular momentum: how to decompose a product of two different angular momenta into a set of single system angular momenta.
- Time-independent perturbation theory: How to approach problems in which the Hamiltonian contains small noncommuting terms.
- Hydrogen atom, including perturbations.
- Connections to classical mechanics: classical limits, symmetries, Hamiltonian formalism, Hamilton-Jacobi equation.

Section 2
Postulates of Quantum Mechanics

Summary

1. The state of a particle is represented by a vector in a Hilbert space.
2. The fundamental state variables x and p of classical mechanics are replaced by Hermitian operators X and P whose matrix elements are well specified in a Hilbert space basis consisting of position eigenstates (states with perfectly defined position x). Any derived dynamical variables $\omega(x, p)$ are replaced by operators Ω defined by the above correspondence.
3. Measurement of any classical variable $\omega(x, p)$ for a quantum state yields only the eigenvalues of the corresponding operator Ω, with the probability of obtaining the eigenvalue ω given by the squared norm of the projection of the state onto the eigenstate corresponding to ω.
4. The state vector evolves according to the Schrödinger equation.

Postulate 1: Representation of Particle States

The state of a particle is represented by a vector $|\psi(t)\rangle$ in a Hilbert space.

What do we mean by this?

We shall define Hilbert space and vectors therein rigorously later; it suffices to say for now that a vector in a Hilbert space is a far more complicated thing than the two numbers x and p that would define the classical state of a particle; the vector is an infinite set of numbers.

The only useful immediate inference we can draw from this statement on its own, based on the definition of *Hilbert space* and *vector*, is that states can be combined linearly. This is interesting, as there is no classical analogue to linear combination of states; for example, if a classical particle only has access to classical state 1 with phase space coordinates (x_1, p_1) and classical state 2 with (x_2, p_2), the particle can only be in one or the other; there is no way to "combine" the two states. Another way of saying this is that quantum mechanics provides for "superposition" of states in a way that classical mechanics does not. But, while interesting, it is not clear what this means or what the experimental implications might be.

Postulate 1: Representation of Particle States (cont.)

The state of a particle is represented by a vector $|\psi(t)\rangle$ in a Hilbert space.

We will see below that Postulates 1 and 3 give rise to the interpretation of the state vector as an object that gives the probability of measuring a particular value for a particular classical observable, depending on what Hilbert space basis the vector is written in terms of. Typically, $|\psi\rangle$ is written in terms of the position basis (a set of Hilbert space vectors with well-defined particle position), in which case $|\psi\rangle$ will give the probability of finding the particle at a given position.

Postulate 2: Correspondence for Classical Variables

The independent variables x and p that describe completely the state of a particle in classical mechanics are represented by Hermitian operators X and P in the Hilbert space of states, with X and P having the following matrix elements when using the position basis for the Hilbert space:

$$\langle x\,|X\,|x'\,\rangle = x\delta\left(x-x'\right) \qquad \langle x\,|P\,|x'\,\rangle = -i\,\hbar\,\frac{d}{dx}\delta\left(x-x'\right) \tag{2.1}$$

We know x and p completely define the classical state of a particle because Newton's Second Law is a second-order differential equation: once x and its first derivative (via p) are specified at an instant in time, all higher-order derivatives are specified.

Postulate 2: Correspondence for Classical Variables (cont.)

The independent variables x and p that describe completely the state of a particle in classical mechanics are represented by Hermitian operators X and P in the Hilbert space of states, with X and P having the following matrix elements when using the position basis for the Hilbert space:

$$\langle x\,|X|x'\,\rangle = x\delta\left(x-x'\right) \qquad \langle x\,|P|x'\,\rangle = -i\,\hbar\,\frac{d}{dx}\delta\left(x-x'\right) \tag{2.2}$$

That is: Pick a basis for the Hilbert space of states that consists of position eigenstates, states that have definite, perfectly defined position. These of course may not be eigenstates of the Hamiltonian and thus may not have definite energy, but we don't care; we don't know about the Hamiltonian yet or the intepretation of its eigenstates. Then, everywhere we see in classical mechanics the position variable x, we replace it with the operator X whose matrix elements are defined as above for any pair of position basis states. This statement is almost a tautology: pick position basis states; then define the X operator such that the position basis states $\{|x\,\rangle\}$ are orthogonal eigenstates of the X operator with eigenvalues $\{x\delta(0)\}$.[a]

[a] We will define and discuss in detail δ functions later.

Postulate 2: Correspondence for Classical Variables (cont.)

The independent variables x and p that describe completely the state of a particle in classical mechanics are represented by Hermitian operators X and P in the Hilbert space of states, with X and P having the following matrix elements when using the position basis for the Hilbert space:

$$\langle x\,|X|x'\,\rangle = x\delta\left(x-x'\right) \qquad \langle x\,|P|x'\,\rangle = -i\,\hbar\,\frac{d}{dx}\delta\left(x-x'\right) \tag{2.3}$$

Why operators? Why are the operators fully specified by matrix elements? Why Hermitian?

We posit that classical variables are replaced by operators because, given the Hilbert space of particle states, the only way to extract real numbers corresponding to classical variables is to assume that there are operators that map from the Hilbert space to itself; such operators are completely specified by their matrix elements between pairs of states in the Hilbert space, and those matrix elements provide the necessary numbers. Why the operators must be Hermitian will be seen in Postulate 3.

Why can we not posit a simpler correspondence, that the operators X and P simply map from the Hilbert space to the real numbers? Because such a framework would just be classical mechanics, for we would be able to assign a specific value of x and p to each state $|\psi\,\rangle$ via $x = X\,|\psi\,\rangle$ and $p = P\,|\psi\,\rangle$.

Postulate 2: Correspondence for Classical Variables (cont.)

Any arbitrary classical dynamical variable $\omega(x, p)$ has a corresponding Hermitian operator

$$\Omega(X, P) = \omega(x \to X, p \to P) \tag{2.4}$$

where we simply replace x and p in ω with X and P to obtain $\Omega(X, P)$.

This is a fairly obvious extension of the first part of this postulate. It is predicated on the fact that any classical variable ω must be a function of x and p because x and p completely define the classical particle state. Since we have above specified a correspondence rule for x and p, this statement carries that rule through to all classical variables.[a]

[a] We shall consider later the complication that arises when ω includes products of x and p; because X and P are non-commuting operators, some thought must be put into how to order X and P in the correspondence.

Postulate 3: Results of Measurements of Classical Variables

Let $\{|\omega\rangle\}$ denote the set of eigenstates of the Hermitian operator with eigenvalues ω. If a particle is in an arbitrary state $|\psi\rangle$, then measurement of the variable corresponding to the operator Ω will yield only the eigenvalues $\{\omega\}$ of Ω. The measurement will yield the particular value ω for that variable with relative probability $P(\omega) = |\langle\omega|\psi\rangle|^2$ and the system will change from state $|\psi\rangle$ to state $|\omega\rangle$ as a result of the measurement being made.

This postulate puts physical meaning to postulates 1 and 2. Those postulates say how we define the particle state and what we replace our classical variables with. This postulate tells us how those operators extract information from the states.

This postulate hinges on the mathematical statement that any valid physical variable operator Ω has eigenstates with eigenvalues. This is just a mathematical result of the assumptions that the states live in a Hilbert space and that the operators Ω must be Hermitian.

At a conceptual level, the postulate means that measurement of a physical quantity *is* the action of the corresponding operator on the state.

Postulate 3: Results of Measurements of Classical Variables (cont.)

Let $\{|\omega\rangle\}$ denote the set of eigenstates of the Hermitian operator with eigenvalues ω. If a particle is in an arbitrary state $|\psi\rangle$, then measurement of the variable corresponding to the operator Ω will yield only the eigenvalues $\{\omega\}$ of Ω. The measurement will yield the particular value ω for that variable with relative probability $P(\omega) = |\langle\omega|\psi\rangle|^2$ and the system will change from state $|\psi\rangle$ to state $|\omega\rangle$ as a result of the measurement being made.

But let's break the statement down carefully:

1. The eigenvalues of Ω are the only values the measured quantity may take on.
2. The measurement outcome is fundamentally probabilistic, and the relative probability[a] of a particular allowed outcome ω is given by finding the projection of $|\psi\rangle$ onto the corresponding eigenstate $|\omega\rangle$. This of course implies that, if $|\psi\rangle$ is an eigenstate of Ω, then the measurement will always yield the corresponding eigenvalue.
3. The measurement process itself changes the state of the particle to the eigenstate $|\omega\rangle$ corresponding to the measurement outcome ω.

---

[a] By *relative* probability, we simply mean that the ratio of the probabilities of two outcomes is given by $P(\omega_1)/P(\omega_2) = |\langle\omega_1|\psi\rangle|^2 \,/\, |\langle\omega_2|\psi\rangle|^2$. The *absolute* probability of a particular outcome requires a normalizing factor that sums over all possible measurement outcomes, to be discussed later.

Postulate 3: Results of Measurements of Classical Variables (cont.)

Let $\{|\omega\rangle\}$ denote the set of eigenstates of the Hermitian operator with eigenvalues ω. If a particle is in an arbitrary state $|\psi\rangle$, then measurement of the variable corresponding to the operator Ω will yield only the eigenvalues $\{\omega\}$ of Ω. The measurement will yield the particular value ω for that variable with relative probability $P(\omega) = |\langle\omega|\psi\rangle|^2$ and the system will change from state $|\psi\rangle$ to state $|\omega\rangle$ as a result of the measurement being made.

The above points are far more than mathematics: they make assumptions about the relationship between physical measurements and the mathematical concepts of eigenstates and eigenvectors.

One could have assumed something simpler: that the measurement outcome is not probabilistic, but is rather the weighted mean of the eigenvalues with $|\langle\omega|\psi\rangle|^2$ providing the weighting factors; and that the act of measurement does not change $|\psi\rangle$. But this would be very similar to classical mechanics.

The assumptions we have chosen to make are the physical content of quantum mechanics and are what distinguish it from classical mechanics.

Postulate 4: Time Evolution of States

The time evolution of the state vector $|\psi(t)\rangle$ is governed by the Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle \tag{2.5}$$

where $H(X, P)$ is the operator obtained from the classical Hamiltonian $\mathcal{H}(x, p)$ via the correspondence $x \longrightarrow X$ and $p \longrightarrow P$.

This statement requires little motivation at a general level: clearly, there must be some time evolution of $|\psi\rangle$ in order for there to be any interesting physics.

There is of course the technical question: why this particular form? One can, to some extent, derive the Schrödinger Equation in various ways, but those methods rely on assumptions, too.[a] Those assumptions may be more intellectually satisfying than simply postulating the Schrödinger Equation, but they provide no definite proof because they simply rely on different assumptions. Experiment provides the ultimate proof that this form is valid.

[a] We shall discuss some of these derivations later when we connect quantum mechanics to classical mechanics at a technical level.

Section 3

Mathematical Preliminaries

Lecture 2:
Linear Vector Spaces, Representations, Linear Independence, Bases

Date Revised: 2008/10/01

Date Given: 2008/10/01

Prologue

We require a fair amount of mathematical machinery to discuss quantum mechanics:

- We must define the space that particle states live in.
- We must define what we mean by the operators that act on those states and give us physical observable quantities.
- We must explore the properties of these operators, primarily those properties that relate to Postulate 3, which says that an operator's eigenvalues are the only physically observable values for the associated physical variable.
- We must understand how states are normalized because of the important relation between the state vector and the relative probabilities of obtaining the spectrum of observable values for a given operator.
- We must also explore the operator analogues of symmetry transformations in classical mechanics; while these do not correspond to physical observables directly, we will see that they are *generated* by physical observables.

Prologue (cont.)

Why so much math? Well, in classical mechanics, we just deal with real numbers and functions of real numbers. You have been working with these objects for many years, have grown accustomed to them, and have good intuition for them. Being Caltech undergrads, calculus is a second language to you. So, in classical mechanics, you could largely rely on your existing mathematical base, with the addition of a few specific ideas like the calculus of variations, symmetry transformations, and tensors.

In QM, our postulates immediately introduce new mathematical concepts that, while having some relation to the 3D real vector space you are familiar with, are significant generalizations thereof. If we taught Hilbert spaces and operators from kindergarten, this would all be second nature to you. But we don't, so you now have to learn all of this math very quickly in order to begin to do QM.

Linear Vector Spaces: Definitions

Let us first discuss the idea of a linear vector space. A linear vector space $\mathbb{V}V$ is a set of objects (called vectors, denoted by $|v\rangle$) and another associated set of objects called scalars (collectively known as a field), along with the following set of rules:

- The vectors have an addition operation (vector addition), $+$, that is closed, meaning that, for any $|v\rangle$ and $|w\rangle$ there exists a $|u\rangle$ in the vector space such that $|u\rangle = |v\rangle + |w\rangle$.

  We may also write the sum as $|v + w\rangle$.

  In defining the set of vectors that make up the vector space, one must also specify how addition works at an algorithmic level: when you add a particular $|v\rangle$ and $|w\rangle$, how do you know what $|u\rangle$ is?
- Vector addition is associative: $(|v\rangle + |w\rangle) + |u\rangle = |v\rangle + (|w\rangle + |u\rangle)$ for all $|u\rangle$, $|v\rangle$, and $|w\rangle$.
- Vector addition is commutative: $|v\rangle + |w\rangle = |w\rangle + |v\rangle$ for any $|v\rangle$ and $|w\rangle$.
- There is a unique vector additive zero or null or identity vector $|0\rangle$: $|v\rangle + |0\rangle = |v\rangle$ for any $|v\rangle$.
- Every vector has a unique vector additive inverse vector: for every $|v\rangle$ there exists a *unique* vector $-|v\rangle$ in the vector space such that $|v\rangle + (-|v\rangle) = |0\rangle$.

Linear Vector Spaces: Definitions (cont.)

- The scalars have an addition operation (scalar addition), $+$, that is closed, so that $a + b$ belongs to the scalar field if a and b do. The addition table must be specified.
- Scalar addition is associative: $a + (b + c) = (a + b) + c$ for any a, b, and c.
- Scalar addition is commutative: $a + b = b + a$ for any a, b.
- A unique scalar additive identity 0 exists: $a + 0 = a$ for any a.
- For any a, a unique scalar additive inverse $-a$ exists with $a + (-a) = 0$.
- The scalars have a multiplication operation (scalar multiplication) that is closed so that the product $a\,b$ belongs to the scalar field if a and b do. The multiplication table must be specified.
- Scalar multiplication is associative, $a\,(b\,c) = (a\,b)\,c$.
- Scalar multiplication is commutative, $a\,b = b\,a$.
- A unique scalar multiplication identity 1 exists: $1\,a = a$ for all a.
- For any $a \neq 0$, a unique scalar multiplicative inverse a^{-1} exists with $a\,a^{-1} = 1$.
- Scalar multiplication is distributive over scalar addition: $a\,(b + c) = a\,b + a\,c$.

Linear Vector Spaces: Definitions (cont.)

- There is a multiplication operation between vectors and scalars (scalar-vector multiplication) that is closed: For any vector $|v\rangle$ and any scalar α, the quantity $\alpha|v\rangle$ is a member of the vector space.

  We may also write the product as $|\alpha v\rangle$.

  Again, one must specify how this multiplication works at an algorithmic level.
- Scalar-vector multiplication is distributive in the obvious way over addition in the vector space: $\alpha\left(|v\rangle + |w\rangle\right) = \alpha|v\rangle + \alpha|w\rangle$
- Scalar-vector multiplication is distributive in the obvious way over addition in the field: $(\alpha + \beta)\,|v\rangle = \alpha|v\rangle + \beta|v\rangle$
- Scalar-vector multiplication is associative in the obvious way over multiplication in the field: $\alpha\left(\beta|v\rangle\right) = (\alpha\beta)\,|v\rangle$

Any vector of the form

$$|u\rangle = \alpha|v\rangle + \beta|w\rangle$$

is called a linear combination and belongs in the space according to the above rules.

Linear Vector Spaces: Definitions (cont.)

Shankar makes fewer assumptions than we do here and states that many of the properties of the scalar field we have assumed can in fact be derived. We choose to assume them because: a) the above assumptions are the standard mathematical definition of a field; and b) if one does not assume the above properties, one has to make some assumptions about how non-trivial the field arithmetic rules are in order to derive them. It's easier, and less prone to criticism by mathematicians, if we do as above rather than as Shankar.

Linear Vector Spaces: Definitions (cont.)

There are a few items that one needs to prove, though:

- The scalar addition identity 0 is consistent with the vector addition identity: $0|v\rangle = |0\rangle$

  Proof: $0|v\rangle + \alpha|v\rangle = (0+\alpha)\,|v\rangle = \alpha|v\rangle$. Since $|0\rangle + \alpha|v\rangle = \alpha|v\rangle$ already, and the identity element is unique, it holds that $0|v\rangle = |0\rangle$.

- Scalar-vector multiplication against the vector addition identity yields the obvious result $\alpha|0\rangle = |0\rangle$

  Proof: $\alpha|0\rangle + \alpha|v\rangle = \alpha\,(|0\rangle + |v\rangle) = \alpha|v\rangle$. Since $|0\rangle + \alpha|v\rangle = \alpha|v\rangle$ already, and the identity element is unique, it holds that $\alpha|0\rangle = |0\rangle$.

- The scalar multiplicative identity is the identity for scalar-vector multiplication also: $1|v\rangle = |v\rangle$

  Proof: $\alpha|v\rangle = (1\alpha)\,|v\rangle = 1\,(\alpha|v\rangle)$; α is arbitrary, so it holds for any $|v\rangle$ that $|v\rangle = 1|v\rangle$

- Vector additive inverses are consistent with scalar additive inverses: $(-1)\,|v\rangle = -|v\rangle$

  Proof: $(-1)\,|v\rangle + |v\rangle = (-1+1)\,|v\rangle = 0|v\rangle = |0\rangle$. Since inverses are unique, it holds that $(-1)\,|v\rangle = -|v\rangle$.

Example 3.1: Real vectors in N spatial dimensions, also known as $\mathbb{R}^N$

You are used to seeing real vectors in 3, and perhaps N, spatial dimensions, defined by the ordered triple

$$|v\rangle \leftrightarrow \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

where the $\{v_j\}$ are all real numbers (the reason for using $\leftrightarrow$ instead of $=$ will become clear in Example 3.4). The vector addition and scalar-vector multiplication algorithms are

$$|v\rangle + |w\rangle \leftrightarrow \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \end{bmatrix}$$

$$\alpha|v\rangle \leftrightarrow \alpha \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} \alpha\, v_1 \\ \alpha\, v_2 \\ \alpha\, v_3 \end{bmatrix}$$

Linear Vector Spaces: Examples (cont.)

Scalar addition and multiplication are just standard addition and multiplication of real numbers. All these operations are closed — *i.e.*, give back elements of the vector space — simply because addition and multiplication of real numbers is closed and because none of the operations change the "triplet" nature of the objects. Extension to N spatial dimensions is obvious. You should carry this example in your head as an intuitive representation of a linear vector space.

Example 3.2: Complex vectors in N spatial dimensions, also known as $\mathbb{C}^N$

Making our first stab at abstraction beyond your experience, let's consider complex vectors in N spatial dimensions. This consists of all ordered N-tuples

$$|v\rangle \leftrightarrow \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$$

where the $\{z_j\}$ are complex numbers, along with the same vector addition and scalar-vector multiplication rules as in the previous example. The space is closed by a logic similar to that used in the real vector space example.

This example is no more complicated than the real vector space $\mathbb{R}^N$. However, your intuition starts to break down because you will no doubt find it hard to visualize even the $N = 2$ example. You can try to imagine it to be something like real 2D space, but now you must allow multiplication by complex coefficients. The next obvious thing is to imagine it to be like real 4D space, but that's impossible to visualize. Moreover, it is misleading because it gives the impression that the space is 4-dimensional, but it really is only two-dimensional. Here is where you must start relying on the math and having only intuitive, not literal, pictures in your head.

Example 3.3: A spin-1/2 particle affixed to the origin

An interesting application of $\mathbb{C}^N$, and one that presents our first example of the somewhat confusing mathematics of quantum mechanics, is spin-1/2 particles such as the electron. As you no doubt learned in prior classes, such particles can be in a state of spin "up" along some spatial direction (say the z axis), spin "down" along that axis, or some linear combination of the two. (Later in the course, we will be rigorous by what we mean about that, but your intuition will suffice for now.) If we fix the particle at the origin so its only degree of freedom is the orientation of its spin axis, then the vector space of states of such particles consists of complex vectors with $N = 2$:

$$|\psi\rangle \leftrightarrow \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

A particle is in a pure spin-up state if $z_2 = 0$ and in a pure spin-down state if $z_1 = 0$.

Linear Vector Spaces: Examples (cont.)

There are many weirdnesses here:

- The particle state can be a linear combination of these two states:

$$|\psi\rangle \leftrightarrow z_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + z_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

  The state is neither perfectly= spin up or perfectly spin down. We will frequently write this state as

$$|\psi\rangle = z_1 |\uparrow\rangle + z_2 |\downarrow\rangle \quad \text{with} \quad |\uparrow\rangle \leftrightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad |\downarrow\rangle \leftrightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

- The particle *lives* in $\mathbb{R}^3$, a real vector space of $N = 3$ dimensions, in that we measure the orientation of its spin axis relative to the axes of that vector space. But the vector space of its quantum mechanical states is $\mathbb{C}^2$, the complex vector space of $N = 2$ dimensions. *The space of QM states is distinct from the space the particle "lives" in!*

Linear Vector Spaces: Examples (cont.)

Example 3.4: The set of all complex-valued functions on a set of discrete points $i\,\frac{L}{n+1}$, $i = 1, \ldots, n$, in the interval $(0, L)$

You are well aware of the idea of a complex function $f(x)$ on an interval $(0, L)$. Here, let's consider a simpler thing, a function an a set of equally spaced, discrete points. The vector $|f\rangle$ corresponding to a particular function is then just

$$|f\rangle \leftrightarrow \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_N) \end{bmatrix} \quad \text{with} \quad x_j = j\,\frac{L}{N+1}$$

You are used to taking linear combinations of functions,

$$h(x) = \alpha\, f(x) + \beta\, g(x)$$

We can do the same thing with these vector objects:

$$|h\rangle = \alpha\,|f\rangle + \beta\,|g\rangle \leftrightarrow \alpha \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_N) \end{bmatrix} + \beta \begin{bmatrix} g(x_1) \\ \vdots \\ g(x_N) \end{bmatrix} = \begin{bmatrix} \alpha\, f(x_1) + \beta\, g(x_1) \\ \vdots \\ \alpha\, f(x_n) + \beta\, g(x_N) \end{bmatrix}$$

Linear Vector Spaces: Examples (cont.)

It is hopefully obvious that this is just a more complicated way of writing $\mathbb{C}^N$: the space is just the set of N-tuples of complex numbers, and, since we can define any function of the $\{x_j\}$ that we want, we can obtain any member of $\mathbb{C}^N$ that we want.

This example lets us introduce the concept of a representation. Given a set of objects and a set of rules for their arithmetic — such as the vector space $\mathbb{C}^N$ — a representation is a way of writing the objects down on paper and expressing the rules. One way of writing down $\mathbb{C}^N$ is simply as the set of all N-tuples of complex numbers. Another way is as the set of all linear combinations of x^a for $a = 1, \ldots, N$ on these discrete points. To give a specific example in $\mathbb{C}^3$:

$$\begin{bmatrix} (1/4)\,L \\ (1/2)\,L \\ (3/4)\,L \end{bmatrix} \leftrightarrow |u\rangle \leftrightarrow x \qquad \begin{bmatrix} (1/16)\,L \\ (1/4)\,L \\ (9/16)\,L \end{bmatrix} \leftrightarrow |v\rangle \leftrightarrow x^2 \qquad \begin{bmatrix} (1/64)\,L \\ (1/8)\,L \\ (27/64)\,L \end{bmatrix} \leftrightarrow |w\rangle \leftrightarrow x^3$$

The vector space elements are $|u\rangle$, $|v\rangle$, and $|w\rangle$. In the column-matrix representation, they are represented by the column matrices. In the functional representation, they are represented by the given functions. We use $\leftrightarrow$ to indicate "represented by" to distinguish it from "equality".

The alert reader will note that the representation in terms of functions is not one-to-one — it is easy to make two functions match up at 3 points but be different elsewhere. We will not worry about this issue now, it will matter later.

Linear Vector Spaces: Examples (cont.)

An aspect of the concept of representation that is confusing is that we usually need to write down a representation to initially define a space. Here, to define $\mathbb{C}^N$, we needed to provide the representation in terms of complex N-tuples. But the space $\mathbb{C}^N$ is more general than this representation, as indicated by the fact that one can write $\mathbb{C}^N$ in terms of the function representation. The space takes on an existence beyond the representation by which it was defined.

In addition to introducing the concept of a representation, this example will become useful as a lead-in to quantum mechanics. You can think of these vectors as the QM wavefunction (something we will define later) for a particle that lives only on these discrete sites $\{x_j\}$. We will eventually take the limit as the spacing $\Delta = \frac{L}{N+1}$ vanishes and N becomes infinite, leaving the length of the interval fixed at L but letting the function now take on a value at any position in the interval $[0, L]$. This will provide the wavefunction for a particle confined in a box of length L.

Finally, one must again be careful not to confuse the space that the particle lives in with the space of its quantum mechanical states. In this case, the former is set of n points on a 1-dimensional line in $\mathbb{R}^1$, while the latter is $\mathbb{C}^N$. When we take the limit $\Delta \to 0$, the particle will then live in the interval $[0, L]$ in $\mathbb{R}^1$, but its space of states will become infinite-dimensional!

Example 3.5: The set of real, antisymmetric $N \times N$ square matrices with the real numbers as the field.

Antisymmetric matrices satisfy $A^T = -A$ where ^{T} indicates matrix transposition. For $N = 3$, these matrices are of the form

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ -a_{12} & 0 & a_{23} \\ -a_{13} & -a_{23} & 0 \end{bmatrix}$$

The vector addition operation is standard matrix addition, element-by-element addition. The scalar arithmetic rules are just addition and multiplication on the real numbers. The scalar-multiplication operation is multiplication of all elements of the matrix by the scalar.

Linear Vector Spaces: Examples (cont.)

It is easy to see that this set satisfies all the vector space rules:

- The sum of two antisymmetric matrices is clearly antisymmetric.
- Addition of matrices is commutative and associative because the element-by-element addition operation is.
- The null vector is the matrix with all zeros.
- The additive inverse is obtained by taking the additive inverse of each element.
- Multiplication of a real, antisymmetric matrix by a real number yields a real, antisymmetric matrix.
- Scalar-vector multiplication is distributive because the a scalar multiplies every element of the matrix one-by-one.
- Scalar-vector multiplication is associative for the same reason.

Note that standard matrix multiplication is *not* included as one of the arithmetic operations here! You can check that the space is not closed under that operation.

Linear Vector Spaces: Examples (cont.)

This example provides a more subtle version of the concept of a representation. There are two aspects to discuss here. First, the example shows that one need not write a vector space in terms of simple column matrices. Here, we use $N \times N$ square matrices instead. The key is whether the objects satisfy the linear vector space rules, not the form in which the objects are written. Second, one can see that this vector space is a representation of $\mathbb{R}^{N(N-1)/2}$: any element has $N(N-1)/2$ real numbers that define it, and the arithmetic rules for matrix addition and scalar multiplication and addition are consistent with the corresponding rules for column-matrix addition and scalar multiplication and addition.

Clearly, one must learn to generalize, to think abstractly beyond a representation of these mathematical objects to the objects themselves. The representation is just what you write down to do calculations, but the rules for the objects are more generic than the representation.

What is the minimal set of vectors needed to construct all the remaining vectors in a vector space? This question brings us to the concepts of linear independence and of a basis for the vector space.

A set of vectors $\{|v_j\rangle\}$ is linearly independent if no one of them can be written in terms of the others. Mathematically: there is no solution to the equation

$$\sum_{j=1}^{n} \alpha_j |v_j\rangle = |0\rangle \tag{3.1}$$

except $\alpha_j = 0$ for all j. The rationale for this definition is straightforward: suppose there were such a set of $\{\alpha_j\}$, and suppose without loss of generality that $\alpha_1 \neq 0$. Then we can rewrite the above as

$$|v_1\rangle = \frac{1}{\alpha_1} \sum_{j=2}^{n} \alpha_j |v_j\rangle \tag{3.2}$$

thereby rewriting $|v_1\rangle$ in terms of the others.

A vector space is defined to have dimension n if the maximal set of linearly independent vectors (excluding $|0\rangle$) that can be found has n members.

Linear Vector Spaces: Linear Independence and Bases (cont.)

We next state two important expansion theorems (The proofs are straightforward, you can look them up in Shankar).

- Given a set of n linearly independent vectors $\{|v_j\rangle\}$ in a n-dimensional vector space, any other vector $|v\rangle$ in the vector space can be expanded in terms of them:

$$|v\rangle = \sum_j \alpha_j |v_j\rangle \tag{3.3}$$

- The above expansion is unique.

Because of the above expansion theorems, any such set of n linearly independent vectors is called a basis for the vector space and is said to span the vector space. The coefficients $\{\alpha_j\}$ for a particular vector $|v\rangle$ are called the components of $|v\rangle$. Equation 3.3 is termed the (linear) expansion of $|v\rangle$ in terms of the basis $\{|v_j\rangle\}$. The vector space is said to be the space spanned by the basis.

Note that, by definition, the concept of linear independence and the linear expansion are representation-independent — both concepts are defined in terms of the vectors and the field elements, not in terms of representations. As usual, you must usually pick a representation to explicitly test for linear independence or to calculate expansion coefficients, but the result must be representation-independent because the definitions are.

Example 3.6: The real and complex vectors on N spatial dimensions

The obvious basis for both of these spaces is

$$|1\rangle \leftrightarrow \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \qquad |2\rangle \leftrightarrow \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \qquad \cdots \qquad |N\rangle \leftrightarrow \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

Other bases are possible, though. For example

$$|1'\rangle \leftrightarrow \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \qquad |2'\rangle \leftrightarrow \begin{bmatrix} 1 \\ -1 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \qquad \cdots \qquad |(N-1)'\rangle \leftrightarrow \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 1 \end{bmatrix} \qquad |N'\rangle \leftrightarrow \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -1 \end{bmatrix}$$

One can prove linear independence by writing down Equation 3.1, giving N equations in the N unknowns $\{\alpha_j\}$ and solving. The first basis just yields the N equations $\alpha_j = 0$ for each j, which implies linear independence. Try the second basis for yourself.

In addition, one can show that $\mathbb{R}^N$ and $\mathbb{C}^N$ are N-dimensional by trying to create a $(N+1)$-dimensional basis. We add to the set an arbitrary vector

$$|v\rangle \leftrightarrow \begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix}$$

where the $\{v_j\}$ are real for $\mathbb{R}^N$ and complex for $\mathbb{C}^N$, and set up Equation 3.1 again. If we use the first basis $\{|j\rangle\}$, one obtains the solution $\alpha_j = v_j$, indicating that any $|v\rangle$ is not linearly independent of the existing set. Since there are N elements of the existing set, the space is N-dimensional.

Note that this proves that $\mathbb{C}^N$, as defined, with a complex field, is N-dimensional, not $2N$-dimensional. If one restricts the field for $\mathbb{C}^N$ to real numbers, then one requires a set of N purely real basis elements and N purely imaginary basis elements, yielding a $2N$-dimensional space. But that is a different space than the one we defined; with a complex field, $\mathbb{C}^N$ is without a doubt N-dimensional.

Example 3.7: Spin-1/2 particle at the origin

We saw in Example 3.3 that this space is just $\mathbb{C}^2$. Here, though, it is useful to get into the physics of different bases. We already stated (without explanation) that the usual orthonormal basis for this space corresponds to spin up and spin down relative to the physical z axis:

$$|\uparrow_z\rangle \leftrightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad |\downarrow_z\rangle \leftrightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Two other reasonable bases are

$$|\uparrow_x\rangle \leftrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad |\downarrow_x\rangle \leftrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$|\uparrow_y\rangle \leftrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} \qquad |\downarrow_y\rangle \leftrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

where $i = \sqrt{-1}$ here.

As the notation suggests, $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$ correspond, respectively, to a particle in a spin up or spin down state relative to the physical x axis, and, similarly, $|\uparrow_y\rangle$ and $|\downarrow_y\rangle$ are the same for the physical y axis. We shall see how these different bases arise as *eigenvectors* of, respectively, the z, x, and y axis spin operators S_z, S_x, and S_y. One can immediately see that, if a particle is in a state of definite spin relative to one axis, it cannot be in a state of definite spin with respect to another — *e.g.*, $|\uparrow_x\rangle = (|\uparrow_z\rangle + |\downarrow_z\rangle)/\sqrt{2}$. This inability to specify spin along multiple axes simultaneously reflects the fact that the corresponding spin operators do not commute, a defining property of quantum mechanics. Much more on this later; certainly, rest assured that this mathematical discussion has significant physical implications.

Following the linear expansion formulae, we can expand the elements of any basis in terms of any other basis; *e.g.*:

$$|\uparrow_y\rangle = \frac{1}{\sqrt{2}}\left[|\uparrow_z\rangle + i\,|\downarrow_z\rangle\right] \qquad |\downarrow_y\rangle = \frac{1}{\sqrt{2}}\left[|\uparrow_z\rangle - i\,|\downarrow_z\rangle\right]$$

$$|\uparrow_z\rangle = \frac{1}{\sqrt{2}}\left[|\uparrow_y\rangle + |\downarrow_y\rangle\right] \qquad |\downarrow_z\rangle = \frac{-i}{\sqrt{2}}\left[|\uparrow_y\rangle - |\downarrow_y\rangle\right]$$

Example 3.8: The set of all complex-valued functions on a set of discrete points $i\,\frac{L}{n+1}$, $i = 1, \ldots, n$, in the interval $(0, L)$, as in Example 3.4

As we have discussed, this space is the same as $\mathbb{C}^N$. The first basis given in the previous example for $\mathbb{C}^N$ is fine and has the advantage of being physically interpreted as having the particle localized at one of the N discrete points: $|i\rangle$ corresponds to the particle being at $x_j = j\,L/(N+1)$. But another basis is the one corresponding to the power law functions x^a, $a = 1, \ldots, N$. For $N = 3$, the representations are

$$x \leftrightarrow |1\rangle \leftrightarrow \begin{bmatrix} (1/4)\,L \\ (1/2)\,L \\ (3/4)\,L \end{bmatrix} \quad x^2 \leftrightarrow |2\rangle \leftrightarrow \begin{bmatrix} (1/16)\,L \\ (1/4)\,L \\ (9/16)\,L \end{bmatrix} \quad x^3 \leftrightarrow |3\rangle \leftrightarrow \begin{bmatrix} (1/64)\,L \\ (1/8)\,L \\ (27/64)\,L \end{bmatrix}$$

If one writes down Equation 3.1, one can show that the only solution is, again, $\alpha_j = 0$ for all i. Let's write the three equations as a matrix equation:

$$\begin{bmatrix} 1/4 & 1/16 & 1/64 \\ 1/2 & 1/4 & 1/8 \\ 3/4 & 9/16 & 27/64 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Recall your linear algebra here: the solution for the $\{\alpha_j\}$ is only nontrivial if the determinant of the matrix vanishes. It does not, so $\alpha_j = 0$ for all i.

Example 3.9: The set of real, antisymmetric $N \times N$ matrices (with the real numbers as the field) as in Example 3.5.

This space is a vector space of dimension $N(N-1)/2$ with one possible basis set just being the real, antisymmetric matrices with two nonzero elements each; for example, for $N = 3$, we have

$$|1\rangle \leftrightarrow \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \quad |2\rangle \leftrightarrow \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad |3\rangle \leftrightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

As with $\mathbb{R}^N$ and $\mathbb{C}^N$, there are many other possible bases. One alternative is

$$|1'\rangle \leftrightarrow \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \quad |2'\rangle \leftrightarrow \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \quad |3'\rangle \leftrightarrow \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$$

One can check that both sets are linearly independent.

Lecture 3:
Inner Product Spaces
Dual Spaces, Dirac Notation, and Adjoints

Date Revised: 2008/10/03

Date Given: 2008/10/03

Inner Product Spaces: Definitions

Now, let us introduce the idea of an inner product, which lets us discuss normalization and orthogonality of vectors.

An inner product is a function that obtains a single complex number from a pair of vectors $|v\rangle$ and $|w\rangle$, is denoted by $\langle v|w\rangle$, and has the following properties:

- positive definiteness: $\langle v|v\rangle \geq 0$ with $\langle v|v\rangle = 0$ only if $|v\rangle = |0\rangle$; *i.e.*, the inner product of any vector with itself is positive unless the vector is the null vector.
- transpose property: $\langle v|w\rangle = \langle w|v\rangle^*$, or changing the order results in complex conjugation.
- linearity: $\langle u|\alpha v + \beta w\rangle = \alpha\langle u|v\rangle + \beta\langle u|w\rangle$

This definition is specific to the case of vector spaces for which the field is the real or complex numbers. Technical problems arise when considering more general fields, and we will only use vector spaces with real or complex fields, so this restriction is not problematic.

Inner Product Spaces: Definitions (cont.)

Some notes:

- It is not necessary to assume $\langle v \,|\, v \rangle$ is real; the transpose property implies it.
- The above also implies antilinearity, $\langle \alpha v + \beta w \,|\, u \rangle = \alpha^* \langle v \,|\, u \rangle + \beta^* \langle w \,|\, u \rangle$
- Inner products are *representation-independent* — the above definitions refer only to the vectors and say nothing about representations. Therefore, if one has two representations of a linear vector space and one wants them to become representations of the same inner product space, the inner product must be defined consistently between the two representations.

Inner Product Spaces: Definitions (cont.)

Now, for some statements of the obvious:

An inner product space is a vector space for which an inner product function is defined.

The length or norm or normalization of a vector $|v\rangle$ is simply $\sqrt{\langle v|v\rangle}$, which we write as $|v|$. A vector is normalized if its norm is 1; such a vector is termed a unit vector. Note that a unit vector can be along any direction; for example, in $\mathbb{R}^3$, you usually think of the unit vectors as being only the vectors of norm 1 along the x, y, and z axes; but, according to our definition, one can have a unit vector along any direction.

The inner product $\langle v|w\rangle$ is sometimes called the projection of $|w\rangle$ onto $|v\rangle$ or vice versa. This derives from the fact that, for $\mathbb{R}^3$, the inner product reduces to

$$\langle v|w\rangle = |v|\,|w|\,\cos\theta_{vw}$$

where θ_{vw} is the angle between the two vectors. In more abstract spaces, it may not be possible to define an angle, but we keep in our minds the intuitive picture from $\mathbb{R}^3$. In general, the two vectors must be normalized in order for this projection to be a meaningful number: when you calculate the projection of a normalized vector onto another normalized vector, the projection is a number whose magnitude is less than or equal to 1 and tells what (quadrature) fraction of $|w\rangle$ lies along $|v\rangle$ and vice versa. We will discuss *projection operators* shortly, which make use of this definition. Note that the term "projection" is not always used in a rigorous fashion, so the context of any discussion of projections is important.

Inner Product Spaces: Definitions (cont.)

Two vectors are orthogonal or perpendicular if their inner product vanishes. This is equivalent to saying that their projections onto each other vanish.

A set of vectors is orthonormal if they are mutually orthogonal and are each normalized; *i.e.*, $\langle v_i | v_j \rangle = \delta_{ij}$ where δ_{ij} is the Kronecker delta symbol, taking on value 1 if $i = j$ and 0 otherwise. We will frequently use the symbol $|i\rangle$ for a member of a set of orthonormalized vectors simply to make the orthonormality easy to remember.

Inner Product Spaces: Definitions (cont.)

Calculating Inner Products

As usual, the above definitions do not tell us algorithmically how to calculate the inner product in any given vector space. The simplest way to do this is to provide the inner products for all pairs of vectors in a particular basis, consistent with the rules defining an inner product space, and to assume linearity and antilinearity. Since all other vectors can be expanded in terms of the basis vectors, the assumptions of linearity and antilinearity make it straightforward to calculate the inner product of any two vectors.

That is, if $\{|j\rangle\}$ is a basis (not necessarily orthonormal), and $|v\rangle = \sum_{j=1}^{n} v_j |j\rangle$ and $|w\rangle = \sum_{j=1}^{n} w_j |j\rangle$, then

$$\langle v|w\rangle = \left\langle \sum_{j=1}^{N} v_j (j) \middle| \sum_{k=1}^{N} w_k (k) \right\rangle$$

where (j) and (k) are the $|j\rangle$ and $|k\rangle$ vectors. Using linearity and antilinearity,

$$\langle v|w\rangle = \sum_{j=1}^{N} v_j^* \left\langle j \middle| \sum_{k=1}^{N} w_k (k) \right\rangle = \sum_{j=1}^{N} \sum_{k=1}^{N} v_j^* w_k \langle j|k\rangle = \sum_{j,k=1}^{N} v_j^* w_k \langle j|k\rangle \qquad (3.4)$$

Once we know all the $\langle j|k\rangle$ inner products, we can calculate the inner product of any two vectors.

Of course, if the basis is orthonormal, this reduces to

$$\langle v \,|w\, \rangle = \sum_{jk} v_j^* w_k \delta_{jk} = \sum_j v_j^* w_j \tag{3.5}$$

and, for an inner product space defined such that component values can only be real numbers, such as $\mathbb{R}^3$ space, we just have the standard dot product. (Note that we drop the full details of the indexing of j and k when it is clear from context.)

With the assumptions that the basis elements satisfy the inner product space rules and of linearity and antilinearity, the transpose property follows trivially. Positive definiteness follows nontrivially from these assumptions for the generic case, trivially for an orthonormal basis.

Note also that there is no issue of representations here — the inner products $\langle j \,|k\, \rangle$ must be defined in a representation-independent way, and the expansion coefficients are representation-independent, so the inner product of any two vectors remains representation-independent as we said it must.

Example 3.10: $\mathbb{R}^N$ and $\mathbb{C}^N$

The inner product for $\mathbb{R}^N$ is the dot product you are familiar with, which happens because the basis in terms of which we first define $\mathbb{R}^N$ is an orthonormal one. The same statement holds for $\mathbb{C}^N$, too, with the complex conjugation of the first member's expansion coefficients. So, explicitly, given two vectors (in $\mathbb{R}^N$ or $\mathbb{C}^N$)

$$|v\rangle \leftrightarrow \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix} \qquad |w\rangle \leftrightarrow \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix}$$

(note the use of arrows instead of equality signs to indicate representation!) their inner product is

$$\langle v|w\rangle = \sum_j v_j^* w_j$$

(Note the equality sign for the inner product, in contrast to the arrows relating the vectors to their representations — again, inner products are representation-independent.) Because the basis is orthonormal, the entire space is guaranteed to satisfy the inner product rules and the spaces are inner product spaces.

Example 3.11: Spin-1/2 particle at the origin

The three bases we gave earlier,

$$|\uparrow_x\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad |\downarrow_x\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$|\uparrow_y\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \end{bmatrix} \qquad |\downarrow_y\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix}$$

$$|\uparrow_z\rangle \leftrightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad |\downarrow_z\rangle \leftrightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

are *each* clearly orthonormal by the algorithm for calculating the $\mathbb{C}^2$ inner product; *e.g.*,

$$\langle\uparrow_y \mid \uparrow_y\rangle = [1\cdot 1 + (-i)\cdot i]/2 = 1 \qquad \langle\uparrow_y \mid \downarrow_y\rangle = [1\cdot 1 + (-i)\cdot(-i)]/2 = 0$$

(Note the complex conjugation of the first element of the inner product!) Hence, according to our earlier argument, the space is an inner product space.

Inner Product Spaces: Examples (cont.)

It is physically interesting to explore the inner products between members of different bases. Some of them are

$$\langle \uparrow_x \mid \uparrow_z \rangle = \frac{1}{\sqrt{2}} \qquad \langle \uparrow_x \mid \downarrow_z \rangle = \frac{1}{\sqrt{2}} \qquad \langle \downarrow_y \mid \uparrow_x \rangle = \frac{1+i}{2} \qquad \langle \downarrow_y \mid \downarrow_x \rangle = \frac{1-i}{2}$$

The nonzero values of the various cross-basis inner products again hint at how definite spin along one direction does not correspond to definite spin along others; *e.g.*, $|\uparrow_x\rangle$ has a nonzero projection onto both $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$.

Inner Product Spaces: Examples (cont.)

Example 3.12: The set of all complex-valued functions on a set of discrete points $i\,\frac{L}{n+1}$, $i = 1, \ldots, n$, in the interval $(0, L)$, as in Example 3.4

We know that this is just a different representation of $\mathbb{C}^N$, but writing out the inner product in terms of functions will be an important lead-in to inner products of QM states on the interval $[0, L]$. Our representation here is

$$|f\rangle \leftrightarrow \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_N) \end{bmatrix} \quad \text{with} \quad x_j = j\,\frac{L}{N+1} \equiv j\,\Delta$$

We use the same orthonormal basis as we do for our usual representation of $\mathbb{C}^N$,

$$|1\rangle \leftrightarrow \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad |2\rangle \leftrightarrow \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \qquad \cdots \qquad |N\rangle \leftrightarrow \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

so that $\langle j\,|k\rangle = \delta_{jk}$.

The inner product of two arbitrary vectors in the space is then

$$\langle f \,|g\, \rangle = \sum_j f^*(x_j)\, g(x_j) \tag{3.6}$$

That is, one multiplies the conjugate of the first function against the second function point-by-point over the interval and sums. The norm of a given vector is

$$\langle f \,|f\, \rangle = \sum_j f^*(x_j)\, f(x_j) = \sum_j |f(x_j)|^2 \tag{3.7}$$

We shall see later how these go over to integrals in the limit $\Delta \longrightarrow 0$.

Example 3.13: The set of real, antisymmetric $N \times N$ matrices (with the real numbers as the field) with conjugation, element-by-element multiplication, and summing as the inner product (*c.f.*, Example 3.5).

Explicitly, the inner product of two elements $|A\rangle$ and $|B\rangle$ is

$$\langle A|B\rangle = \sum_{jk} A^*_{jk} B_{jk} \tag{3.8}$$

where jk indicates the element in the jth row and kth column. We include the complex conjugation for the sake of generality, though in this specific example it is irrelevant. Does this inner product satisfy the desired properties?

- Positive definiteness: yes, because the inner product squares away any negative signs, resulting in a positive sum unless all elements vanish.
- Transpose: yes, because the matrix elements are real and real multiplication is commutative.
- Linearity: yes, because the expression is linear in B_{kl}.

Inner Product Spaces: Examples (cont.)

This inner product makes this space a representation of $\mathbb{R}^{N(N-1)/2}$ as an inner product space. Let's write down normalized versions of the bases we considered previously:

$$|1\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \qquad |2\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad |3\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$|1'\rangle \leftrightarrow \frac{1}{2}\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \qquad |2'\rangle \leftrightarrow \frac{1}{2}\begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \qquad |3'\rangle \leftrightarrow \frac{1}{2}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$$

It is fairly obvious that the first basis is an orthogonal basis. By direct calculation, you can quickly see that the second basis is *not* orthogonal.

As a digression, we note that the inner product can also be written as

$$\langle A\,|B\rangle = \sum_{jk} A^*_{jk} B_{jk} = \sum_{jk} A^{T*}_{kj} B_{jk} = \mathrm{Tr}(A^\dagger B)$$

$$\text{where} \qquad M^\dagger_{jk} = M^*_{kj} \qquad \text{and} \qquad \mathrm{Tr}(M) = \sum_j M_{jj} \qquad \text{for any matrix } M$$

Here, we begin to see where matrix multiplication can become useful in this vector space. But note that it only becomes useful as a way to calculate the inner product.

Dual Spaces and Dirac Notation

We have seen examples of representing vectors $|v\rangle$ as column matrices for $\mathbb{R}^N$ and $\mathbb{C}^N$. This kind of column matrix representation is valid for any linear vector space because the space of column matrices, with standard column-matrix addition and scalar-column-matrix multiplication and scalar addition and multiplication, is itself a linear vector space. Essentially, column matrices are just a bookkeeping tool for keeping track of the coefficients of the basis elements.

Inner Product Spaces: Dual Spaces, Dirac Notation, and Adjoints (cont.)

When we begin to consider inner product spaces, we are naturally led to the question of how the inner product works in this column-matrix representation. We immediately see

$$\begin{aligned}\langle v\,|w\,\rangle &= \sum_{j,k=1}^{N} v_j^*\, w_k \langle j\,|k\,\rangle \\ &= \begin{bmatrix} v_1^* & \cdots & v_N^* \end{bmatrix} \begin{bmatrix} \langle 1\,|1\,\rangle & \cdots & \langle 1\,|N\,\rangle \\ \vdots & \ddots & \vdots \\ \langle N\,|1\,\rangle & \cdots & \langle N\,|N\,\rangle \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_N \end{bmatrix}\end{aligned} \tag{3.9}$$

That is, there is an obvious matrix representation of the inner product operation. When the basis is orthonormal, the above simplifies to

$$\langle v\,|w\,\rangle = \sum_{j=1}^{N} v_j^*\, w_j = \begin{bmatrix} v_1^* & \cdots & v_N^* \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_N \end{bmatrix} \tag{3.10}$$

Inner Product Spaces: Dual Spaces, Dirac Notation, and Adjoints (cont.)

Purely for calculational and notational convenience, the above equation for the orthonormal basis case leads us to define, for any vector space $\mathbb{V}$, a partner space, called the dual space $\mathbb{V}^*$, via a row-matrix representation. That is, for a vector $|v\rangle$ in $\mathbb{V}$ with its standard column-matrix representation

$$|v\rangle \leftrightarrow \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \tag{3.11}$$

we **define** a dual vector $\langle v|$ in the dual space $\mathbb{V}^*$ by its *row-matrix representation*

$$\langle v| \leftrightarrow \begin{bmatrix} v_1^* & v_2^* & \cdots & v_n^* \end{bmatrix} \tag{3.12}$$

A key point is that the dual space $\mathbb{V}^*$ is not identical to the vector space $\mathbb{V}$ and is not a vector space because the rules for scalar-vector multiplication are different: since there is a complex conjugation in the definition of the row-matrix representation, $\langle \alpha\, v| = \langle v|\, \alpha^*$ holds. (The placement of the α^* makes no difference to the meaning of the expression; we place the α^* after $\langle v|$ for reasons to be discussed soon.)

Inner Product Spaces: Dual Spaces, Dirac Notation, and Adjoints (cont.)

(Though $\mathbb{V}^*$ is not a vector space, we might consider simply defining a dual vector space to be a set that satisfies all the vector space rules except for the complex conjugation during scalar-vector multiplication. It would be a distinct, but similar, mathematical object.)

Those with strong mathematical backgrounds may not recognize the above definition. The standard definition of the dual space $\mathbb{V}^*$ is the set of all linear functions from $\mathbb{V}$ to its scalar field; *i.e.*, all functions on $\mathbb{V}$ that, given an element $|v\rangle$ of $\mathbb{V}$, return a member α of the scalar field associated with $\mathbb{V}$. These functions are also called *linear functionals*, *linear forms*, *one-forms*, or *covectors*. We shall see below why this definition is equivalent to ours for the cases we will consider. If you are not already aware of this more standard definition of dual space, you may safely ignore this point!

Inner Product Spaces: Dual Spaces, Dirac Notation, and Adjoints (cont.)

With the definition of the dual space, and assuming we have the expansions $|v\rangle = \sum_{j=1}^{N} v_j |j\rangle$ and $|w\rangle = \sum_{j=1}^{N} w_j |j\rangle$ in terms of an orthonormal basis for $\mathbb{V}$, we may now see that the inner product $\langle v|w\rangle$ can be written as the matrix product of the row-matrix representation of the dual vector $\langle v|$ and the column-matrix representation of the vector $|w\rangle$:

$$\langle v|w\rangle = \sum_j v_j^* w_j = \begin{bmatrix} v_1^* & v_2^* & \cdots & v_n^* \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \langle v||w\rangle \qquad (3.13)$$

Again, remember that the representations of $\langle v|$ and $|w\rangle$ in terms of matrices are how our initial definitions of them are made, and are convenient for calculational purposes, but the representations are just that, representations; they are not the same thing as $\langle v|$ and $|w\rangle$, the latter have a more abstract existence.

Inner Product Spaces: Dual Spaces, Dirac Notation, and Adjoints (cont.)

The above now explains our comment about the dual space being the space of linear functionals: there is a one-to-one correspondence between the $\langle v\,|$ dual vectors and the linear functionals $\langle v\,|\;\rangle$ that accept a vector $|w\rangle$ and returns a number α by taking the inner product. In fact, one can show that any linear functional mapping from $\mathbb{V}$ to the field can be decomposed in terms of inner-product operations $\langle v\,|\;\rangle$. Mathematicians use the linear functional definition because it is more generic and connects to other concepts; for example, one-forms more easily generalize to vector spaces with *curvature*, which we will most definitely not discuss in this course, and are connected to the differential curl operator. The most likely place you will encounter such objects are in a course in general relativity. I'll bet, though that, like me, most of you can live without appreciating this subtlety the first time through...

It is standard practice to denote the vector $|v\rangle$ belonging to $\mathbb{V}$ as a ket and the dual vector $\langle v\,|$ belonging to $\mathbb{V}^*$ as a bra. These definitions are termed Dirac notation. Depending on the circumstances, we will use the dual space, the Dirac notation, or both naming schemes.

Inner Product Spaces: Dual Spaces, Dirac Notation, and Adjoints (cont.)

We note that the basis vectors and their corresponding dual vectors satisfy

$$|j\rangle \leftrightarrow \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \langle j| \leftrightarrow \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix} \tag{3.14}$$

where each is matrix is nonzero only in its jth element. The above lets us write

$$|v\rangle = \sum_j v_j |j\rangle = \sum_j \langle j|v\rangle |j\rangle \qquad \langle v| = \sum_j \langle j| v_j^* = \sum_j \langle j|\langle v|j\rangle \tag{3.15}$$

where $v_j = \langle j|v\rangle$ and $v_j^* = \langle v|j\rangle$ simply follow from the expansion of $|v\rangle$ in terms of $\{|j\rangle\}$.

Inner Product Spaces: Dual Spaces, Dirac Notation, and Adjoints (cont.)

Example 3.14: Spin-1/2 particle at the origin

Let's list the matrix representations of some vectors and their dual vectors (kets and bras) for the sake of being explicit:

$$|\uparrow_x\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \langle\uparrow_x| \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \end{bmatrix}$$

$$|\downarrow_x\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad \langle\downarrow_x| \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \end{bmatrix}$$

$$|\uparrow_y\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \end{bmatrix} \qquad \langle\uparrow_y| \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -i \end{bmatrix}$$

$$|\downarrow_y\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix} \qquad \langle\downarrow_y| \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \end{bmatrix}$$

$$|\uparrow_z\rangle \leftrightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \langle\uparrow_z| \leftrightarrow \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$$|\downarrow_z\rangle \leftrightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \langle\downarrow_z| \leftrightarrow \begin{bmatrix} 0 & 1 \end{bmatrix}$$

Inner Product Spaces: Dual Spaces, Dirac Notation, and Adjoints (cont.)

We can check the same inner products we did before, this time evaluating the inner products using the matrix representation (Equation 3.13) rather than representation-free sum over products of coefficients (Equation 3.5):

$$\langle \uparrow_y || \uparrow_y \rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -i \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} = \frac{1}{2}(1+1) = 1 = \langle \uparrow_y \mid \uparrow_y \rangle$$

$$\langle \downarrow_y || \uparrow_y \rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} = \frac{1}{2}(1-1) = 0 = \langle \downarrow_y \mid \uparrow_y \rangle$$

$$\langle \downarrow_y || \uparrow_x \rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{2}(1+i) = \langle \downarrow_y \mid \uparrow_x \rangle$$

Inner Product Spaces: Dual Spaces, Dirac Notation, and Adjoints (cont.)

We can now derive the linear expansions we wrote in Example 3.7: we use Equation 3.15 along with evaluation of the inner products using the matrix representation; *e.g.*,

$$
\begin{aligned}
|\downarrow_z\rangle &= \sum_j \langle j \,|\, \downarrow_z \rangle |j\rangle = \langle \uparrow_y \,|\, \downarrow_z \rangle |\uparrow_y\rangle + \langle \downarrow_y \,|\, \downarrow_z \rangle |\downarrow_y\rangle \\
&= \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -i \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) |\uparrow_y\rangle + \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) |\downarrow_y\rangle \\
&= \frac{-i}{\sqrt{2}} \left[|\uparrow_y\rangle - |\downarrow_y\rangle \right]
\end{aligned}
$$

This is a good example of a situation in which one has to avoid getting confused about what should be written out in matrix representation and what should not. The inner products $\langle \uparrow_y \,|\, \downarrow_z \rangle$ and $\langle \downarrow_y \,|\, \downarrow_z \rangle$ are written as matrix products. We could replace $|\uparrow_y\rangle$ and $|\downarrow_y\rangle$ by their matrix representations also. But keep in mind two things: 1) if you replace the vectors on the right side of the equation by column matrix representation, you must do the same on the left side, too: vectors and their representations are not the same thing; 2) the matrices making up the inner product do not act on the column matrix representation of the vectors by matrix multiplication, as indicated by the parentheses in the expression; the scalar result of the inner product multiplies the column matrices for the vectors.

Example 3.15: The set of real, antisymmetric $N \times N$ matrices, as in Example 3.5.

This is a particularly interesting example because you have to confront the many representations a particular inner product space can have. Let's consider the orthonormal basis we wrote down for the $N = 3$ case in Example 3.9:

$$|1\rangle \leftrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \quad |2\rangle \leftrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad |3\rangle \leftrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

Now, let's construct two new elements of the space via linear combinations; you will recognize these as the $|1'\rangle$ and $|2'\rangle$ normalized but non-orthogonal elements we previously constructed:

$$|1'\rangle = \frac{1}{\sqrt{2}}(|2\rangle + |3\rangle) \leftrightarrow \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$|2'\rangle = \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle) \leftrightarrow \frac{1}{2} \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

Inner Product Spaces: Dual Spaces, Dirac Notation, and Adjoints (cont.)

Let's consider three ways of taking the inner product $\langle 1'|2'\rangle$. The first is the obvious way, using the explicit definition we had for the inner product for this space in terms of the defining representations:

$$\begin{aligned}\langle 1'|2'\rangle &= \sum_{j,k=1}^{3} (1')_{jk}(2')_{jk} \\ &= \frac{1}{4}\left[0\cdot 0 + 1\cdot 1 + 0\cdot 1 + (-1)\cdot(-1) + 0\cdot 0 + 1\cdot 0 + 0\cdot(-1) + (-1)\cdot 0 + 0\cdot 0\right] \\ &= \frac{1}{2}\end{aligned}$$

This above makes use of the representation we used to define the space, but makes no use of the generic column- and row-matrix representations we developed for an arbitrary inner product space with an orthonormal basis.

Inner Product Spaces: Dual Spaces, Dirac Notation, and Adjoints (cont.)

Now, let's use the representation-free sum over coefficients, Equation 3.5, which makes use of the orthonormality of the basis but not the matrix representation of the vector and dual vector space:

$$\langle 1' | 2' \rangle = \sum_{j=1}^{3} (1')_j (2')_j = \frac{1}{\sqrt{2}} (0 \cdot 1 + 1 \cdot 1 + 1 \cdot 0) = \frac{1}{2}$$

Finally, let's write it out in terms of a matrix product of the matrix representations of the vector and dual vector spaces:

$$\langle 1' | 2' \rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{2}$$

Thus, we see that there are two *different* matrix representations of this space: the one we used to define the space, and the one that appears when an orthonormal basis is used for the space. These are different in that they don't look the same; but they are the same in that all the operations we have defined in one representation can be carried over to the other and vice versa in a consistent fashion. Clearly, though, it can be confusing to represent the same abstract object $|v\rangle$ in two different ways — as an $N \times N$ real, antisymmetric matrix and as a $N(N-1)/2$-element column matrix — but this a key concept you must become accustomed to.

Inner Product Spaces: Dual Spaces, Dirac Notation, and Adjoints (cont.)

Representations as a Tool for Simplification and Unification

Hopefully, the previous two examples have illustrated the value of the matrix representation of inner product spaces once an orthonormal basis has been established — once you have established an orthonormal basis and expanded all the elements of the space in terms of that basis, you know the column-matrix representation of any element of the inner product space (and the row-matrix representation for its dual vector) and it is, frequently, arithmetically easier to take inner products using the matrix representation than to use the defining rule for the inner product in the space. Essentially, by recognizing the simpler underlying structure present once an orthonormal basis has been defined, we simplify the operations we must do on the space.

Another point is that the use of matrix representations allows us to unify different spaces, to realize that they are the same in spite of the apparent differences in the way they are defined. Mathematically, this is termed an isomorphism; the spaces are said to be isomorphic. In particular, any inner product space of dimension N with a real (complex) field looks like $\mathbb{R}^N$ ($\mathbb{C}^N$) as far as any of the properties of the inner product space go. Of course, once one introduces additional operations on some spaces, this isomorphism may not be carried through to those operations. But the idea of isomorphism and the isomorphism of inner product spaces corresponding to different physical objects will be a theme we will return to repeatedly in this course.

Inner Product Spaces: Dual Spaces, Dirac Notation, and Adjoints (cont.)

Adjoints

We **define** the process of converting from vector to dual vector (ket to bra) and vice versa as taking the adjoint. $\langle v |$ is the adjoint of $|v\rangle$ and vice versa. In terms of the orthonormal basis matrix representation, there is a simple algorithm for this: complex conjugate and transpose.

From the above definition, the properties of complex numbers, and the definition of inner product, we can derive rules for taking the adjoint of any combination of bras, kets, and scalar coefficients:

- scalar coefficients: When one encounters a bra or ket with a scalar coefficient, the scalar coefficient must be complex conjugated (in addition to taking the adjoint of the bra or ket); *i.e.*, the adjoint of $\alpha|v\rangle$ is $\langle v|\alpha^*$ and vice versa. Again, the placement of α^* on the right is purely notational.

Inner Product Spaces: Dual Spaces, Dirac Notation, and Adjoints (cont.)

- inner products: To determine the adjoint of an inner product $\langle v | w \rangle$, we use the fact that the inner product is just a complex number and so taking the adjoint of the inner product just corresponds to complex conjugation. But we know from the definition of inner product that complex conjugation of an inner product corresponds to exchanging the positions of the two vectors, so we see that the adjoint of $\langle v | w \rangle$ is $\langle w | v \rangle$. Thus, when we encounter **inner** products of bras and kets, we take the adjoint by simply reversing the order and converting bras to kets and vice versa, consistent with our rule for bras and kets alone with the addition of order-reversal. The need for order reversal is why we place scalar coefficients of bras to their right; the notation is now consistent.
- sums: the adjoint of a sum is just the sum of the adjoints because complex conjugation and matrix transposition both behave this way.
- products: Suppose one has an arbitrary product of inner products, scalar coefficents, and a bra or ket. (There can be nothing more complicated because the result would not be a bra or ket and hence could not be in the vector space or the dual vector space.) Our rules above simply imply that one should reverse the order of all the elements and turn all bras into kets and vice versa, even the ones in inner products. That is, for the ket

$$|u\rangle = \alpha_1 \cdots \alpha_k \langle w_1 | v \rangle \cdots \langle w_m | v_m \rangle | v \rangle \tag{3.16}$$

where the $\{\alpha_j\}$, $\{|w_j\rangle\}$, and $\{|v_j\rangle\}$ are arbitrary (the index matchups mean nothing), we have that the adjoint is

$$\langle u| = \langle v|\langle v_1|w_1\rangle \cdots \langle v_m|w_m\rangle \alpha_1^* \cdots \alpha_k^* \tag{3.17}$$

- vector and dual vector expansions: We may write our vector and dual vector expansions as

$$|v\rangle = \sum_j \langle j|v\rangle |j\rangle = \sum_j |j\rangle\langle j|v\rangle \qquad \langle v| = \sum_j \langle v|j\rangle\langle j| \tag{3.18}$$

where we have simply exchanged the order of the inner product and the bra or ket; this is fine because the inner product is just a scalar. We see that the above expansions are fully consistent with our rules for taking adjoints of sums and products.

Lecture 4:
Inner Product Theorems
Gram-Schmidt Orthogonalization
Subspaces

Date Revised: 2008/10/06
Date Given: 2008/10/06

Inner Product Theorems

The definition of inner product immediately gives rise to some useful results. We will only give the basic idea of the proofs; you can find the details in Shankar.

"Law of Cosines"

$$|v + w|^2 = |v|^2 + |w|^2 + 2\,\mathcal{R}\left(\langle v\,|w\,\rangle\right) \tag{3.19}$$

where $\mathcal{R}(z)$ is the real part of the complex number z. This is proven by simply expanding out the inner product implied by the left side. The relation is named as it is because it reduces to the law of cosines when $|v\,\rangle$ and $|w\,\rangle$ belong to $\mathbb{R}^3$.

The astute reader will see that the sign on the inner product term is different than what one usually sees in the law of cosines, $c^2 = a^2 + b^2 - 2\,a\,b\,\cos\gamma$. This is because the angle γ is not the angle between $\vec{a}$ and $\vec{b}$ at the origin; it is the supplementary angle to that angle, hence the sign flip on the cos term. The following diagram this explicitly.

Inner Product Spaces: Theorems (cont.)

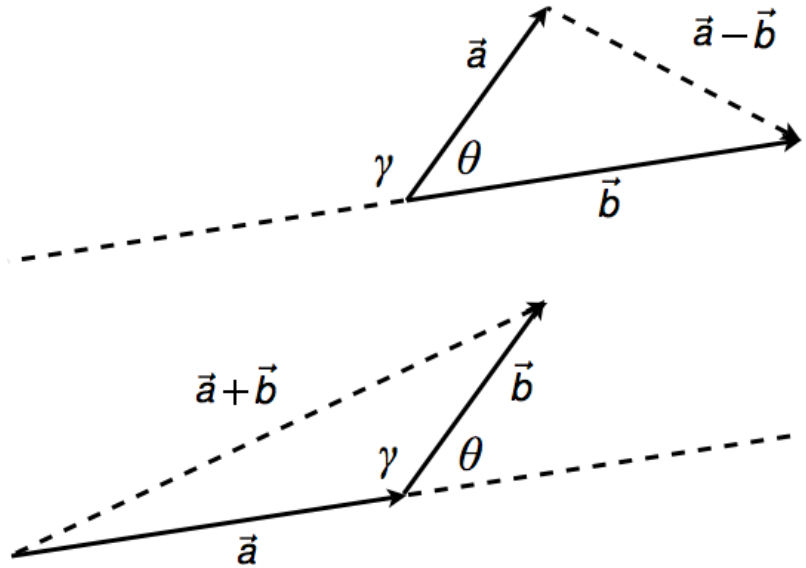


Diagram to illustrate sign of last term in law of cosines in $\mathbb{R}^2$. The law of cosines conventionally involves the angle γ because that is the angle opposite to the vector $\vec{a}+\vec{b}$. The inner product of two vectors gives the cosine of the angle between them when they are placed at the origin, θ. There is a sign flip between the two cosines because they are supplementary angles (sum to π radians). We want to use the θ angle instead of the γ angle in the generic form of the law of cosines, hence the sign flip.

Schwarz Inequality

$$|\langle v\,|w\,\rangle|^2 \leq |v|^2|w|^2 \tag{3.20}$$

The Schwarz inequality simply states that the inner product of two vectors can be no larger than the product of their lengths. More simply, if one divides out the norms of the vectors, it states that the inner product of two unit vectors can be no greater in magnitude than 1. When we think about the interpretation of the inner product as the projection of one vector onto the other, this makes sense; the projection of a unit vector onto another unit vector can be no larger than 1.

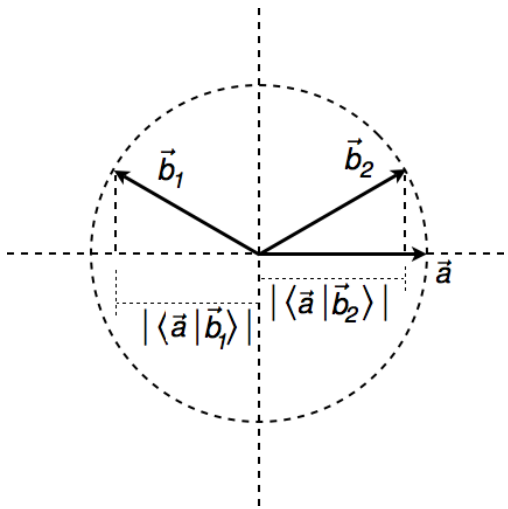


Diagram illustrating Schwarz inequality in $\mathbb{R}^2$ for vectors of unit length. All three vectors $\vec{a}$, $\vec{b}_1$, and $\vec{b}_2$ have unit length (as indicated by the fact that they all end on the circle). Clearly, the projections $|\langle\vec{a}\,|\vec{b}_1\,\rangle|$ and $|\langle\vec{a}\,|\vec{b}_2\,\rangle|$ are both less than 1. Note how the sign of the projection does not affect the result.

The inequality is proven by applying the law of cosines to the vector

$$|v'\rangle = |v\rangle - \langle w|v\rangle \frac{|w\rangle}{|w|^2} = |v\rangle - \langle \widehat{w}|v\rangle |\widehat{w}\rangle \tag{3.21}$$

where $|\widehat{w}\rangle = \widehat{w}/|w|$ is the unit vector along the $|w\rangle$ direction, and using the positive definiteness of the norm of $|v'\rangle$.

Clearly, $|v'\rangle$ is the piece of $|v\rangle$ that is orthogonal to $|w\rangle$. The Schwarz inequality devolves to an equality if $|v\rangle = \lambda|w\rangle$ for some λ; *i.e.*, if $|v\rangle$ and $|w\rangle$ are the same up to a multiplicative constant, indicating they point in the same (or opposite) direction.

Note that the above vector may also be written as

$$|v'\rangle = |v\rangle - |\widehat{w}\rangle\langle\widehat{w}|v\rangle \tag{3.22}$$

We see the expression $|\widehat{w}\rangle\langle\widehat{w}|$ as we did when writing out bras and kets as sums of components along the vectors of an orthonormal basis. Such objects we will see are **projection operators** because they project out the part of the vector they operate on along the unit vector comprising the operator.

Triangle Inequality

$$|v + w| \leq |v| + |w| \tag{3.23}$$

This is a direct result of the law of cosines, arising from the fact that $2\mathcal{R}(\langle v|w\rangle) \leq 2|\langle v|w\rangle| \leq 2|v||w|$. The inequality devolves to an equality only if $|v\rangle = \lambda|w\rangle$ with λ real and positive.

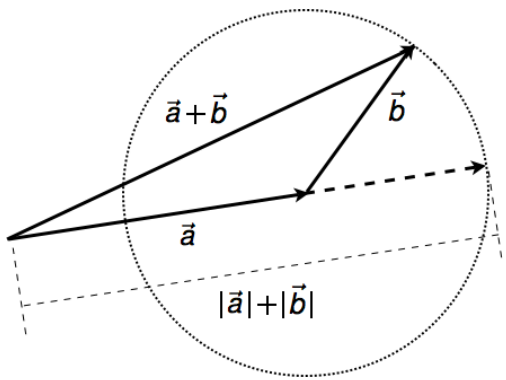


Diagram illustrating triangle inequality in $\mathbb{R}^2$, where it expresses the fact that the length of the sum of two vectors can be no larger than the sum of their individual lengths, and equality occurs when the vectors are coaligned. The circle is centered on the start of $\vec{b}$ and has radius equal to $|\vec{b}|$, so indicates the locus of possible endpoints of $\vec{a}+\vec{b}$; one particular example is given for the orientation of $\vec{b}$ and $\vec{a}+\vec{b}$. The dashed line indicates the maximum length possibility, with $\vec{a}$ and $\vec{b}$ coaligned so that $|\vec{a}+\vec{b}| = |\vec{a}| + |\vec{b}|$.

Gram-Schmidt Orthogonalization

Given n linearly independent vectors $\{|v_j\rangle\}$, one can construct from them an orthonormal set of the same size.

The basic idea is to orthogonalize the set by subtracting from the jth vector the projections of that vector onto the j-1 previous vectors, which have already been orthogonalized. Dirac notation makes the projection operations more obvious.

We begin by using the first ket from the original set, creating a normalized version:

$$|1'\rangle = |v_1\rangle \qquad |1\rangle = \frac{|1'\rangle}{\sqrt{\langle 1'|1'\rangle}} \tag{3.24}$$

Then, the second member of the orthogonal and orthonormal sets are

$$|2'\rangle = |v_2\rangle - \frac{|1'\rangle\langle 1'|v_2\rangle}{\langle 1'|1'\rangle} = |v_2\rangle - |1\rangle\langle 1|v_2\rangle \qquad |2\rangle = \frac{|2'\rangle}{\sqrt{\langle 2'|2'\rangle}} \tag{3.25}$$

and so on; the generic formula is

$$|j'\rangle = |v_j\rangle - \sum_{k=1}^{j-1} \frac{|k'\rangle\langle k'|v_k\rangle}{\langle k'|k'\rangle} = |v_j\rangle - \sum_{k=1}^{j-1} |k\rangle\langle k|v_k\rangle \qquad |j\rangle = \frac{|j'\rangle}{\sqrt{\langle j'|j'\rangle}} \tag{3.26}$$

Inner Product Spaces: Gram-Schmidt Orthogonalization (cont.)

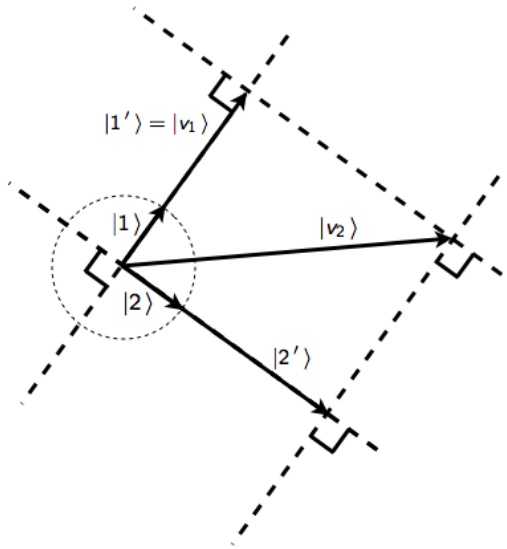


Diagram illustrating Gram-Schmidt orthogonalization in $\mathbb{R}^2$. The first vector $|v_1\rangle$ is simply normalized to obtain $|1\rangle$ but otherwise left unchanged. We subtract off from the second vector $|v_2\rangle$ the projection along $|1\rangle$, leaving $|2'\rangle$. We then normalize $|2'\rangle$ to obtain $|2\rangle$. $|1\rangle$ and $|2\rangle$ are clearly orthogonal and normalized. (The circle has unity radius.)

One proves this theorem inductively, showing that if the first j-1 vectors have been orthogonalized, then the jth vector created via the above formula is orthogonal to the first j-1. The $|j\rangle$ are manifestly normalized.

Gram-Schmidt orthogonalization lets us conclude what we intuitively expect: for an inner product space of dimension n, we can construct an orthonormal basis for the space from any other basis. Shankar proves this point rigorously, but it is easy to see intuitively: the Gram-Schmidt procedure tells us that any linearly independent set of n vectors yields a mutually orthogonal set of n vectors, and it is fairly obvious that a mutually orthogonal set of n vectors is linearly independent. If the initial set of LI vectors matches the dimension of the space, then the new orthonormal set is a basis for the space.

Subspaces

It almost goes without saying that a subspace of a linear vector space or inner product space is a subset of the space that is itself a vector or inner product space. Since the subset inherits the algebraic operations (and inner product, if one exists) from the parent space, the only substantive requirement is that the subspace be closed under the vector addition and scalar-vector multiplication operations. One can show that the parent space's null vector must be in any subspace, and hence there is always one element of overlap between any two subspaces.

Given two subspaces $\mathbb{V}_1$ and $\mathbb{V}_2$ of a vector space $\mathbb{V}$, the sum or direct sum of the two subspaces, denoted by $\mathbb{V}_1 \oplus \mathbb{V}_2$, is the set of all linear combinations of vectors in $\mathbb{V}_1$ and $\mathbb{V}_2$. Note that, since $\mathbb{V}_1$ and $\mathbb{V}_2$ are subspaces of some larger vector space $\mathbb{V}$, it is already known that one may add vectors from $\mathbb{V}_1$ and $\mathbb{V}_2$ together.

Note that $\mathbb{V}_1 \oplus \mathbb{V}_2$ is not the same as $\mathbb{V}_1 \cup \mathbb{V}_2$. $\mathbb{V}_1 \oplus \mathbb{V}_2$ consists of *all* linear combinations of the form

$$|v\rangle = \alpha_1 |v_1\rangle + \alpha_2 |v_2\rangle \tag{3.27}$$

where $|v_1\rangle$ is in $\mathbb{V}_1$ and $|v_2\rangle$ is in $\mathbb{V}_2$. When both α_1 and α_2 are nonzero, $|v\rangle$ belongs to neither $\mathbb{V}_1$ nor $\mathbb{V}_2$, but lives in the part of $\mathbb{V}_1 \oplus \mathbb{V}_2$ outside $\mathbb{V}_1$ and $\mathbb{V}_2$. On the other hand, $\mathbb{V}_1 \cup \mathbb{V}_2$ consists only of the linear combinations for which at least one of α_1 and α_2 vanish (*i.e.*, the trivial linear combination in which no combining is done!). The following diagram may help to illustrate the distinctions.

Subspaces (cont.)

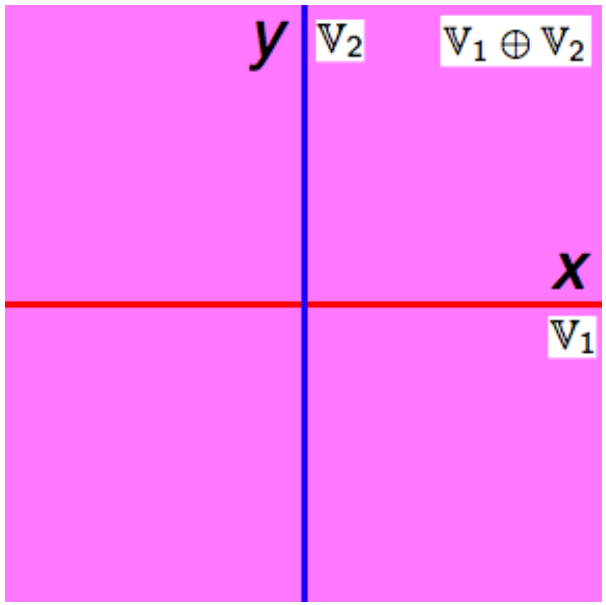


Diagram illustrating subspaces in $\mathbb{R}^2$. The first subspace $\mathbb{V}_1$ consists of all vectors along the x-axis only, indicated by the red line. The second subspace $\mathbb{V}_2$ consists of vectors along the y-axis only, indicated by the blue line. The union $\mathbb{V}_1 \cup \mathbb{V}_2$ consists of all vectors either along the red line or the blue line. The direct sum $\mathbb{V}_1 \oplus \mathbb{V}_2$ consists of all linear combinations of vectors along the red line or the blue line, so covers the entire plane, indicated by the pink shading. $\mathbb{V}_1 \oplus \mathbb{V}_2$ is much bigger than $\mathbb{V}_1$, $\mathbb{V}_2$, or $\mathbb{V}_1 \cup \mathbb{V}_2$.

The most trivial subspace an inner product space $\mathbb{V}$ can have is the set of all vectors of the form $\alpha|v\rangle$ where $|v\rangle$ is some element in $\mathbb{V}$: these are just all the vectors along $|v\rangle$. Given a basis $\{|v_j\rangle\}$, the entire space $\mathbb{V}$ is just the direct sum of the subspaces of this type for each basis element. That is, if we define $\mathbb{V}_j = \{\alpha|v_j\rangle\}$, the set of all scalar multiples of the jth basis element, then

$$\mathbb{V} = \mathbb{V}_1 \oplus \mathbb{V}_2 \oplus \cdots \oplus \mathbb{V}_N$$

Example 3.16: $\mathbb{R}^N$ and $\mathbf{C}^N$

Each of these have a variety of subspaces. Each can be viewed as being the N-element direct sum of their $N = 1$ versions (à la what we just said):

$$\mathbb{R}^N = \overbrace{\mathbb{R} \oplus \mathbb{R} \cdots \oplus \mathbb{R}}^{N \text{ terms}} \qquad \mathbb{C}^N = \overbrace{\mathbb{C} \oplus \mathbb{C} \cdots \oplus \mathbb{C}}^{N \text{ terms}}$$

or, perhaps a direct sum of many copies of $\mathbb{R}^2$ with one $\mathbb{R}$ thrown in if N is odd, or direct sums of $\mathbb{R}^M$ of various M, etc., etc.

Example 3.17: The set of all complex-valued functions on a set of discrete points $i\,\frac{L}{n+1}$, $i = 1, \ldots, n$, in the interval $(0, L)$, as in Example 3.4

The analogue of the above when considering the function representation would be to select functions that are zero at various points. For example, if $N = 4$, the functions that are always zero on x_1 and x_2 are one subspace, the functions that always vanish on x_3 and x_4 are a different subspace, and the full space is the direct sum of these two subspaces. Each subspace is isomorphic to a function space on two discrete points, and hence to $\mathbb{C}^2$.

Example 3.18: Spin-1/2 particle at the origin

This is just $\mathbb{C}^2$ and so the subspaces are fairly boring, but the physical interpretation is interesting. One can consider the two subspaces $\{\alpha|\uparrow_z\rangle\}$ and $\{\alpha|\downarrow_z\rangle\}$ where α is any complex number: these subspaces consist of either spin up or spin down states only. (Note that, even if you restrict to $|\alpha| = 1$, there are still an infinity of elements in each subspace because α is complex.) But one could alternately consider the subspaces $\{\alpha|\uparrow_x\rangle\}$ and $\{\alpha|\downarrow_x\rangle\}$ or $\{\alpha|\uparrow_y\rangle\}$ and $\{\alpha|\downarrow_y\rangle\}$. One recovers the full space by direct sum of the two subspaces in each circumstance, but these provide some alternate subspaces of $\mathbb{C}^2$.

Example 3.19: Real antisymmetric, imaginary symmetric, and anti-Hermitian matrices.

The aforementioned set of real, antisymmetric $N \times N$ matrices with a real number field form a subspace of the set of anti-Hermitian matrices[a], also with a real field. The set of purely imaginary $N \times N$ symmetric matrices with a real field are also a subspace of the anti-Hermitian matrices. The direct sum of the real antisymmetric matrices and the purely imaginary symmetric matrices gives the entire space of anti-Hermitian matrices. Specifically, the three groups are

$$A_R = \begin{bmatrix} 0 & a_{12} & a_{13} \\ -a_{12} & 0 & a_{23} \\ -a_{13} & -a_{23} & 0 \end{bmatrix} \qquad A_I = \begin{bmatrix} 0 & i\,b_{12} & i\,b_{13} \\ i\,b_{12} & 0 & i\,b_{23} \\ i\,b_{13} & i\,b_{23} & 0 \end{bmatrix}$$

$$A_A = \begin{bmatrix} 0 & a_{12} + i\,b_{12} & a_{13} + i\,b_{13} \\ -a_{12} + i\,b_{12} & 0 & a_{23} + i\,b_{23} \\ -a_{13} + i\,b_{13} & -a_{23} + i\,b_{23} & 0 \end{bmatrix}$$

[a] complex matrices A for which $(A^*)^T = -A$ where * is element-by-element complex conjugation and ^{T} is matrix transposition.

Subspaces (cont.)

Let's check a number of things:

- Real antisymmetric matrices are a subset of anti-Hermitian matrices because they do not change under complex conjugation and they pick up a sign under transposition. Similarly, purely imaginary symmetric matrices are a subset because they pick up a sign under complex conjugation but do not change under transposition.
- We have already shown that real antisymmetric matrices are closed. Purely imaginary matrices symmetric matrices are closed under addition and multiplication by real numbers because neither operation can change the fact they are purely imaginary or symmetric.
- We have already shown that real antisymmetric matrices are an inner product space. Purely imaginary symmetric matrices are also an inner product space because the complex conjugation in the inner-product formula ensures positive definiteness. The transpose and linearity rules are also satisfied.
- Any sum of a real antisymmetric matrix and a purely imaginary symmetric matrix is immediately anti-Hermitian because the real part of the sum is guaranteed to change sign and the imaginary part to keep its sign under transposition. Any anti-Hermitian matrix can be decomposed in terms of a real antisymmetric and purely imaginary symmetric matrix simply by breaking it element-by-element into real and imaginary parts.

Lecture 5:
Linear Operators

Date Revised: 2008/10/13
Date Given: 2008/10/08

Prologue

Now that we have defined inner product spaces, we have largely completed the work of defining the space that the states of a physical system live in. This is not enough, as physical states are not static. To make measurements and to obtain the dynamics of the system, we need *operators* that transform states into other states. According to postulates 2, 3, and 4, operators tell us how to carry classical mechanics over to quantum mechanics, how measurements work and how they affect states, and how to time-evolve states.

Linear Operators (cont.)

An operator Ω transforms a vector into a (possibly the same) vector, $\Omega|v\rangle = |w\rangle$ and transforms a dual vector into a (also possibly the same) dual vector, $\langle v|\Omega = \langle u|$.

Note that the action of the operator on $\langle v|$ is not necessarily the bra corresponding to the operation of the operator on $|v\rangle$; *i.e.*, $\langle \Omega v| \neq \langle v|\Omega$ in general (though it will be true in some cases). This is why we stated above $\langle v|\Omega = \langle u|$ instead of $\langle v|\Omega = \langle w|$.

Linear Operators (cont.)

However, it is also true that, once the operation of Ω on vectors has been set, there is no freedom in the action of Ω on dual vectors. One sees this as follows: suppose

$$\Omega|v_1\rangle = |w_1\rangle \qquad \Omega|v_2\rangle = |w_2\rangle \qquad \langle v_1|\Omega = \langle u_1| \qquad \langle v_2|\Omega = \langle u_2|$$

Then we have the 4 independent relations

$$\langle v_1|w_1\rangle = \langle v_1|\Omega|v_1\rangle = \langle u_1|v_1\rangle \qquad \langle v_1|w_2\rangle = \langle v_1|\Omega|v_2\rangle = \langle u_1|v_2\rangle$$
$$\langle v_2|w_1\rangle = \langle v_2|\Omega|v_1\rangle = \langle u_2|v_1\rangle \qquad \langle v_2|w_2\rangle = \langle v_1|\Omega|v_1\rangle = \langle u_2|v_2\rangle$$

More generally, given N linearly independent vectors in a N-dimensional vector space, there will be N^2 relations of the above type. Specifying the action of Ω on these N vectors requires $2\,N^2$ numbers (the expansion coefficients of the $\{|u_j\rangle\}$ and $\{|w_j\rangle\}$). The N^2 expansion coefficients of the $\{|w_j\rangle\}$ were set when the action of Ω on the vector space was defined. The N^2 relations thus determine the N^2 expansion coefficients of the $\{|u_j\rangle\}$ from those of the $\{|w_j\rangle\}$. Thus, to determine the action of Ω on any N-element linearly independent set $\{\langle v_j|\}$, one need only look at Ω's action on the corresponding $\{|v_j\rangle\}$. Finally, if the action of Ω on the full set of vectors in the vector space is specified, then the action of Ω on the full set of dual vectors is specified by just picking linearly independent sets of the above type for each $|v\rangle$ and working out the relations.

Linear Operators (cont.)

We specialize to linear operators, those satisfying typical linearity relations:

$$\Omega\left(\alpha\left|v\right\rangle+\beta\left|w\right\rangle\right)=\alpha\,\Omega\left|v\right\rangle+\beta\,\Omega\left|w\right\rangle \qquad \left(\left\langle v\right|\alpha+\left\langle w\right|\beta\right)\Omega=\alpha\left\langle v\right|\Omega+\beta\left\langle w\right|\Omega \tag{3.28}$$

Such operators are convenient, of course, because their action is completely specified by their action on the vector space's basis vectors, which we shall come to momentarily. We specialize to linear operators because it is the simplest possible choice and it has been verified that quantum mechanics using only linear operators matches experimental predictions.

Our argument about the relation between the action of Ω on vectors and dual vectors simplifies now: once the action of Ω on an orthonormal basis $\{\left|j\right\rangle\}$ has been specified, then our argument indicates that this specifies its action on the orthonormal basis $\{\left\langle j\right|\}$. For a linear operator, specifying its action on an orthonormal basis then gives the action on the entire space by linearity, so the full action of Ω on all vectors and dual vectors is specified.

Example 3.20: Identity Operator

That's an easy one: for any $|v\rangle$, it returns $|v\rangle$:

$$I|v\rangle = |v\rangle$$

The only thing to point out here is that there is not just one identity operator; there is an identity operator for each vector space.

Example 3.21: Projection Operators

Given an inner product space $\mathbb{V}$ and a subspace $\mathbb{V}_P$, a projection operator is the operator that maps a vector $|v\rangle$ into its projection onto that subspace. An equivalent definition is: a projection operator is any operator of the form

$$P = \sum_j |v_j\rangle\langle v_j| \tag{3.29}$$

where the $\{|v_j\rangle\}$ are members of an orthonormal basis (not necessarily all the members!). That is, each term calculates the inner product of the vector $|v\rangle$ it acts on with the unit vector $|v_j\rangle$, then multiplies the result by $|v_j\rangle$ to recover a vector instead of a number. One can see that the two definitions are equivalent by recognizing that the subspace $\mathbb{V}_P$ is just the space spanned by the set $\{|v_j\rangle\}$; alternatively, given the subspace $\mathbb{V}_P$, one should pick the $\{|v_j\rangle\}$ to be any orthonormal basis for $\mathbb{V}_P$.

Linear Operators (cont.)

It is important for the vectors to be an orthonormal set in order to really pick out the piece of $|v\rangle$ in the subspace. To give a trivial counterexample, consider the vector $|v\rangle = v|j\rangle$, where $|j\rangle$ is an orthonormal basis element, and the projection operator $P = |v\rangle\langle v|$. Clearly, the output vector is always a vector in the subspace $\mathbb{V}_j$ spanned by $|j\rangle$ because $|v\rangle$ is in that subspace. But, let's act on $|v\rangle$ with P:

$$P|v\rangle = |v\rangle\langle v|v\rangle = |v|^2|v\rangle = |v|^2 v|j\rangle$$

Since the original projection operator was not composed of normalized vectors, the normalization of the result is funny: it is not the projection of $|v\rangle$ onto the $\mathbb{V}_j$ subspace, but rather $|v|^2$ times that projection.

The above is just a normalization problem. A larger problem arises when one considers a projection operator composed of two non-orthonormal vectors. For example, in $\mathbb{R}^2$, consider the projection operator

$$P = |\widehat{x}\rangle\langle\widehat{x}| + \left|\frac{\widehat{x}+\widehat{y}}{\sqrt{2}}\right\rangle\left\langle\frac{\widehat{x}+\widehat{y}}{\sqrt{2}}\right|$$

where $|\widehat{x}\rangle$ and $|\widehat{y}\rangle$ are the x and y unit vectors. The vectors used to construct P are normalized but not orthogonal. The subspace spanned by the vectors making up the operator is the entire space, $\mathbb{R}^2$, because the two vectors are linearly independent and the space is already known to be 2-dimensional. Let's try acting on $|\widehat{y}\rangle$ (using antilinearity of the bra):

$$P|\widehat{y}\rangle = |\widehat{x}\rangle\langle\widehat{x}|\widehat{y}\rangle + \left|\frac{\widehat{x}+\widehat{y}}{\sqrt{2}}\right\rangle\frac{1}{\sqrt{2}}\left(\langle\widehat{x}|\widehat{y}\rangle + \langle\widehat{y}|\widehat{y}\rangle\right) = \frac{1}{\sqrt{2}}\left|\frac{\widehat{x}+\widehat{y}}{\sqrt{2}}\right\rangle$$

Since the subspace spanned by the vectors making up the operator is the space, this projection operator ought to have returned $|\widehat{y}\rangle$; it did not.

Linear Operators (cont.)

We can explicitly see that, if the projection operator is composed of an orthonormal set, it does indeed recover the portion of $|v\rangle$ in the subspace spanned by that set. Let's consider an orthonormal basis $\{|j\rangle\}$ of a n-dimensional inner product space, and let's consider a projection operator onto the subspace $\mathbb{V}_{1\ldots m}$ spanned by the first m of the basis elements. (We can always reorder the basis elements so the ones that we want to use are the first m.) That projection operator is

$$P = \sum_{j=1}^{m} |j\rangle\langle j|$$

Acting on an arbitrary $|v\rangle$ in the space (with expansion $\sum_{j=1}^{n} v_j|j\rangle$) with P thus gives

$$P|v\rangle = \left(\sum_{j=1}^{m} |j\rangle\langle j|\right)\left(\sum_{k=1}^{n} v_k|k\rangle\right) = \sum_{j=1}^{m}\sum_{k=1}^{n} v_k|j\rangle\langle j|k\rangle = \sum_{j=1}^{m}\sum_{k=1}^{n} v_k|j\rangle\delta_{jk}$$
$$= \sum_{j=1}^{m} v_j|j\rangle$$

which is, by our original expansion of $|v\rangle$, the piece of $|j\rangle$ in the subspace spanned by the first m of the $\{|j\rangle\}$.

Linear Operators (cont.)

It is frequently useful to rewrite the identity operator in terms of projection operators:

$$I = \sum_{j=1}^{n} |j\rangle\langle j| \tag{3.30}$$

That this sum is indeed the identity operator can be seen by using the same proof we just made above but taking $m = n$. Then the result is

$$\left[\sum_{j=1}^{n} |j\rangle\langle j|\right] |v\rangle = \left[\sum_{j=1}^{n} |j\rangle\langle j|\right] \sum_{k=1}^{n} v_k |k\rangle = \sum_{j=1}^{n} v_j |j\rangle = |v\rangle$$

It follows from the above that a projection operator P that projects onto the subspace $\mathbb{V}_P$ is the identity operator I_P for $\mathbb{V}_P$. When $\mathbb{V}_P = \mathbb{V}$, one recovers I for the full space $\mathbb{V}$.

A final note: frequently, we will use the subscript $_j$ to denote projection on the subspace spanned by the single orthonormal basis element $|j\rangle$:

$$P_j = |j\rangle\langle j|$$

Of course, P_j is not specified until you specify a basis, so the meaning of P_j will always depend on context. But this is standard notation.

Example 3.22: Derivative Operator on Function Spaces on Discrete Points

Returning to the Example 3.4, let's create something that looks like taking the derivative. Let's use the orthonormal basis $\{|j\rangle\}$ consisting of the functions that are 1 at x_j and 0 elsewhere (*i.e.*, just like the standard basis of $\mathbb{C}^N$). Then define

$$D_R|j\rangle = -\frac{|j\rangle - |j-1\rangle}{\Delta} \quad \text{for} \quad j \neq 1 \qquad D_R|1\rangle = -\frac{|1\rangle - |N\rangle}{\Delta} \qquad x_j = j\,\frac{L}{N+1} = j\,\Delta$$

Then we have

$$\begin{aligned} D_R|f\rangle = D_R \sum_{j=1}^{N} f(x_j)|j\rangle &= -\frac{1}{\Delta}\left[f(x_1)\left(|1\rangle - |N\rangle\right) + \sum_{j=2}^{N} f(x_j)\left(|j\rangle - |j-1\rangle\right)\right] \\ &= \left[\sum_{j=1}^{N-1} \frac{f(x_{j+1}) - f(x_j)}{\Delta}|j\rangle\right] + \frac{f(x_1) - f(x_N)}{\Delta}|N\rangle \end{aligned}$$

The output function looks like the right-going discrete derivative of $f(x_j)$! We looped around the end at the last point; that would not be necessary in the limit $\Delta \longrightarrow 0$. You are well aware from calculus that taking the derivative of continuous functions is a linear operation, the same holds here.

Example 3.23: Spin-1/2 Operators

Let's give our first example of an operator that returns an observable, the spin projection of spin-1/2 particle states. We simply take as definitions

$$S_z|\uparrow_z\rangle = \frac{\hbar}{2}|\uparrow_z\rangle \qquad S_z|\downarrow_z\rangle = -\frac{\hbar}{2}|\downarrow_z\rangle \tag{3.31}$$

$$S_x|\uparrow_x\rangle = \frac{\hbar}{2}|\uparrow_x\rangle \qquad S_x|\downarrow_x\rangle = -\frac{\hbar}{2}|\downarrow_x\rangle \tag{3.32}$$

$$S_y|\uparrow_y\rangle = \frac{\hbar}{2}|\uparrow_y\rangle \qquad S_y|\downarrow_y\rangle = -\frac{\hbar}{2}|\downarrow_y\rangle \tag{3.33}$$

We have really put the cart before the horse here because we never explained why the states that we defined as $|\uparrow_z\rangle$, $|\downarrow_z\rangle$, $|\uparrow_x\rangle$, etc. corresponded to physical states with spin projection along $+z$, $-z$, $+x$, etc. But neglecting that motivational problem, which we will deal with later, it is clear that the above is a perfectly valid definition of a set of operators, and they now have some physical meaning: these operators tell us the spin projection along particular axes of particular states.

We have only specified the spin operators in terms of the states they leave unchanged (up to normalization). These states are a complete basis for the space in each case, so this is sufficient. But let us look at how they change other states. For example, using the above and some results derived in Example 3.14:

$$S_y|\downarrow_z\rangle = S_y \frac{-i}{\sqrt{2}}\left(|\uparrow_y\rangle - |\downarrow_y\rangle\right) = -\frac{i\,\hbar}{2\sqrt{2}}\left(|\uparrow_y\rangle + |\downarrow_y\rangle\right) = -\frac{i\,\hbar}{2}|\uparrow_z\rangle$$

That is, S_y converts $|\downarrow_z\rangle$ to $|\uparrow_z\rangle$, modulo a normalization factor. We will make use of the "raising" behavior later. For now, it simply serves to show that, having defined the action of S_y on the $|\uparrow_y\rangle$ and $|\downarrow_y\rangle$ basis states, we can now calculate the action on any state, a point that we will state more generally next. The same holds for S_x and S_z.

Example 3.24: Rotation Operators in $\mathbb{R}^3$.

Another example, discussed by Shankar, is that of rotation operators in $\mathbb{R}^3$. Read this example.

Linear Operators (cont.)

Linear Operator Action on Basis Vectors, Matrix Elements, Matrix Representation of Linear Operators

The main advantage to staying with linear operators is that their action on any vector is defined purely by their action on a set of basis vectors. Given a set of basis vectors $\{|j\rangle\}$ and a vector $|j\rangle = \sum_j v_j|j\rangle$ expanded in terms of them, we may write

$$\Omega|v\rangle = \Omega\sum_j v_j|j\rangle = \sum_j v_j\Omega|j\rangle \tag{3.34}$$

It is useful to break $\Omega|j\rangle$ into components by rewriting the expansion using the identity operator written out using projection operators, $I = \sum_k |k\rangle\langle k|$:

$$\Omega|v\rangle = \sum_j v_j\Omega|j\rangle = \sum_j v_j \sum_k |k\rangle\langle k|\Omega|j\rangle$$

We define the projection of $\Omega|j\rangle$ onto $|k\rangle$ as Ω_{kj}, $\Omega_{kj} = \langle k|\Omega|j\rangle$. These are just numbers. The expression can then be rewritten

$$\Omega|v\rangle = \sum_{jk} \Omega_{kj}v_j|k\rangle = \sum_{jk} |k\rangle\Omega_{kj}v_j$$

thereby giving the components of the result along the various $\{|k\rangle\}$.

The above expression looks like matrix multiplication of a $n \times n$ matrix against a single-column n-row matrix:

$$[\Omega|v\rangle]_k = \langle k|\Omega|v\rangle = \sum_j \Omega_{kj} v_j \tag{3.35}$$

This makes sense: we were able to represent our vectors via single-column matrices (kets) and our dual vectors as single-row matrices (bras); it is consistent for operators to be represented as $n \times n$ matrices (where n is the dimensionality of the vector space) and the kj element (kth row, jth column) is just the projection of the action of Ω on $|j\rangle$ onto $|k\rangle$. We have thus found the matrix representation of the operator Ω in the column-matrix representation of the vector space with orthonormal basis $\{|j\rangle\}$.

Linear Operators (cont.)

We may of course derive similar relations for the operation of Ω on a bra:

$$\langle v\,|\Omega = \sum_j \langle j\,|\Omega v_j^* = \sum_{jk} v_j^* \langle j\,|\Omega|k\,\rangle\langle k\,| = \sum_{jk} v_j^* \Omega_{jk} \langle k\,|$$

$$\text{or} \quad [\langle v\,|\Omega]_k = \sum_j v_j^* \Omega_{jk} \tag{3.36}$$

which again looks like matrix multiplication, this time of a $n \times n$ matrix on a single-row, n-column matrix on its left. Ω_{jk} is the projection of the action of Ω on $\langle j\,|$ onto $\langle k\,|$ (note the transposition of the indices relative to the ket case). The matrix representation of Ω is thus also consistent with the row-matrix representation of the dual vector space with orthonormal basis $\{\langle j\,|\}$.

We note that the above relation corroborates the statement we made at the start of our discussion of operators that specifying the action of Ω on a linear basis for $\mathbb{V}$ also full determines its action on a linear basis for $\mathbb{V}^*$. Here, the Ω_{kj} are the N^2 numbers that give the action of Ω on any ket $|v\,\rangle$, as indicated in Equation 3.35. But these same N^2 numbers appear in Equation 3.36, which expresses the action of Ω on any bra $\langle v\,|$.

Let us summarize our point about matrix representations of operators: given a linear operator Ω on a n-dimensional inner product space $\mathbb{V}$ with an orthonormal basis $\{|j\rangle\}$ (and corresponding dual space with orthonormal basis $\{\langle j|\}$), we may write a $n \times n$ matrix representation of the operator Ω with elements Ω_{kj} given by

$$\Omega_{kj} = \langle k|\Omega|j\rangle \tag{3.37}$$

and the action of this matrix on the column-matrix representation of $\mathbb{V}$ and the row-matrix representation of $\mathbb{V}^*$ is consistent with the operation of Ω on the elements of $\mathbb{V}$ and $\mathbb{V}^*$. Matrix representations of operators will be the tool we use to do much of quantum mechanics.

Example 3.25: Projection Operators Revisited

Projection operators have a very simply matrix representation:

$$(P_j)_{kl} = \langle k\,|P_j|m\,\rangle = \langle k\,|j\,\rangle\langle j\,|m\,\rangle = \delta_{kj}\delta_{jm} \tag{3.38}$$

That is, P_j is an empty matrix except for a 1 in the jj element. Conveniently, this extends the consistency of the matrix representation scheme for bras and kets if we define an outer product between vectors,

$$|v\,\rangle\langle w\,| = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \begin{bmatrix} w_1^* & w_2^* & \cdots & w_n^* \end{bmatrix} = \begin{bmatrix} v_1 w_1^* & v_1 w_2^* & \cdots & v_1 w_n^* \\ v_2 w_1^* & v_2 w_2^* & \cdots & v_2 w_n^* \\ \vdots & & \ddots & \vdots \\ v_n w_1^* & v_n w_2^* & \cdots & v_n w_n^* \end{bmatrix} \tag{3.39}$$

$$\text{or} \quad [|v\,\rangle\langle w\,|]_{km} = v_k w_m^* \tag{3.40}$$

because, for a projection operator $P_j = |j\,\rangle\langle j\,|$, we have $v_k = \delta_{kj}$ and $w_m = \delta_{jm}$ as shown earlier.

Linear Operators (cont.)

For $\mathbb{R}^3$ with the standard x, y, z basis for the column-matrix representation, the projection operators that project onto the x, y, and z axes are

$$P_x \leftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad P_y \leftrightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad P_z \leftrightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The projection operators into various planes are

$$P_{xy} \leftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad P_{yz} \leftrightarrow \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad P_{xz} \leftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Linear Operators (cont.)

We note that one can write a projection operator in a matrix representation corresponding to a basis that does not match the natural basis of the projection operator. For example, in the above basis, the operator that projects onto the subspace defined by the vector $(\widehat{x}+\widehat{y})/\sqrt{2}$ is

$$P_{\frac{\widehat{x}+\widehat{y}}{\sqrt{2}}} \leftrightarrow \left(\frac{|\widehat{x}\rangle + |\widehat{y}\rangle}{\sqrt{2}}\right)\left(\frac{\langle\widehat{x}| + \langle\widehat{y}|}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1 & 0\end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix}1 & 1 & 0\\1 & 1 & 0\\0 & 0 & 0\end{bmatrix}$$

The key is to use the outer product form, writing the column- and row-matrix representations of the bras and kets making up the outer product.

Example 3.26: Derivative Operator on Function Spaces on Discrete Points, as in Example 3.22

We defined the action of the derivative operator on the space of functions on N discrete points as follows:

$$D_R|j\rangle = \begin{cases} -\left(|1\rangle - |N\rangle\right)/\Delta & j = 1 \\ -\left(|j\rangle - |j-1\rangle\right)/\Delta & j \neq 1 \end{cases} \qquad x_j = j\,\frac{L}{N+1} = j\,\Delta$$

It's easy to calculate the matrix elements:

$$(D_R)_{kj} = \langle k|D_R|j\rangle = \begin{cases} \left(\delta_{k,N} - \delta_{k,1}\right)/\Delta & j = 1 \\ \left(\delta_{k,j-1} - \delta_{k,j}\right)/\Delta & j \neq 1 \end{cases}$$

Let's check that we get the expected result from this matrix representation for $N = 4$:

$$D_R|f\rangle \leftrightarrow \frac{1}{\Delta}\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} f(x_1) \\ f(x_2) \\ f(x_3) \\ f(x_4) \end{bmatrix} = \begin{bmatrix} [f(x_2) - f(x_1)]/\Delta \\ [f(x_3) - f(x_2)]/\Delta \\ [f(x_4) - f(x_3)]/\Delta \\ [f(x_1) - f(x_4)]/\Delta \end{bmatrix}$$

You have to be careful about what is meant by $\delta_{k,j-1}$. For example, for row 1 of the matrix, $k = 1$, we have $\delta_{k,j-1} = 1$ when $j - 1 = k = 1$, so $j = 2$ has the nonzero element.

Example 3.27: Spin-1/2 Operators

Let's write out the matrix representations of the S_x, S_y, and S_z operators in the orthonormal basis $\{|\uparrow_z\rangle, |\downarrow_z\rangle\}$. We state without derivation that they are

$$S_z \leftrightarrow \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad S_x \leftrightarrow \frac{\hbar}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad S_y \leftrightarrow \frac{\hbar}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{3.41}$$

Checking S_z is easy because we defined it by its action on $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$, which are the elements of the orthonormal basis we are using for this matrix representation:

$$S_z|\uparrow_z\rangle \leftrightarrow \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{\hbar}{2}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \leftrightarrow \frac{\hbar}{2}|\uparrow_z\rangle$$

$$S_z|\downarrow_z\rangle \leftrightarrow \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{\hbar}{2}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \leftrightarrow \frac{\hbar}{2}|\downarrow_z\rangle$$

which reproduce Equation 3.31.

We have defined S_x and S_y in terms of their action on $|\uparrow_x\rangle$, $|\downarrow_x\rangle$ and $|\uparrow_y\rangle$, $|\downarrow_y\rangle$, respectively, so one must apply the matrix representations of the operators in this basis to the matrix representations of those vectors in this basis. The latter were given in Example 3.14. Let's try one example here:

$$S_y|\downarrow_y\rangle \leftrightarrow \frac{\hbar}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix} = -\frac{\hbar}{2}\begin{bmatrix} 1 \\ -i \end{bmatrix} \leftrightarrow -\frac{\hbar}{2}|\downarrow_y\rangle$$

which matches Equation 3.33.

Lecture 6:
Linear Operators Continued
Date Revised: 2008/10/13
Date Given: 2008/10/10

Bilinear Form

With the concept of matrix elements, we show a useful way of writing out operators, the bilinear form. (This is not explicitly discussed in Shankar.) The idea is to generalize the way we write projection operators in terms of outer products to write all operators in a similar fashion. Let $\{|j\rangle\}$ be an orthonormal basis for our space. Then, inserting the projection version of the identity operator on both sides of an operator Ω, we have

$$\Omega = I\,\Omega\,I = \sum_{j,k=1}^{n} |k\rangle\langle k|\Omega|j\rangle\langle j| = \sum_{j,k=1}^{n} |k\rangle\Omega_{kj}\langle j| = \sum_{j,k=1}^{n} \Omega_{kj}|k\rangle\langle j| \tag{3.42}$$

That's it: the idea is that, once you have specified an orthonormal basis, you can write any operator as a bilinear expression in the basis kets and bras, with the coefficients simply being all the matrix elements in that basis. Note that this form depends on your choice of basis because the $\{|j\rangle\}$ and the $\{\Omega_{kj}\}$ depend on that choice. That said, the above statement is an absolute equality, not a representation equivalence (*i.e.*, not just "$\leftrightarrow$"). But note that this particular form will only be useful if one is working with expansions of the vectors and dual vectors in terms of these specific $\{|j\rangle\}$ and $\{\langle j|\}$.

Matrix Elements of Products of Operators

Using projection operators, it is easy to show that the matrix elements of a product of two operators is found by simple matrix multiplication of the matrices representing the two operators:

$$[\Omega\Lambda]_{jk} = \langle j\,|\Omega\Lambda|k\,\rangle = \langle j\,|\Omega\left(\sum_m |m\,\rangle\langle m\,|\right)\Lambda|k\,\rangle = \sum_m \langle j\,|\Omega|m\,\rangle\langle m\,|\Lambda|k\,\rangle$$

$$= \sum_m \Omega_{jm}\Lambda_{mk} \tag{3.43}$$

Our matrix representation scheme remains consistent (operator product = standard matrix product).

Operator Adjoints

Since operators are completely determined by their actions on vectors, there should be a way to define the adjoint of an operator that is consistent with our vector adjoint definition. We in fact make the definition to require consistency: the adjoint $\Omega^\dagger$ of an operator Ω yields the adjoint of the vector $\Omega|v\rangle$ when it acts on $\langle v|$:

$$\text{if} \quad |w\rangle = \Omega|v\rangle \quad \text{then} \quad \langle w| = \langle v|\Omega^\dagger \tag{3.44}$$

defines $\Omega^\dagger$. Recall our discussion showing that specifying the action of Ω on all vectors $|v\rangle$ fully determines its action on all dual vectors $\langle v|$; the converse also holds, so defining $\Omega^\dagger$ by its action on dual vectors thereby defines its action on vectors. For linear operators, we can obtain an algorithmic formula for obtaining $\Omega^\dagger$ from Ω via any matrix representation:

$$\left[\Omega^\dagger\right]_{jk} = \langle j|\Omega^\dagger|k\rangle = [\Omega|j\rangle]^\dagger\,|k\rangle = \left(\langle k|\,[\Omega|j\rangle]\right)^* = \left(\langle k|\Omega|j\rangle\right)^* = \Omega^*_{kj} \tag{3.45}$$

where we reversed the order of the inner product because, at that point, $[\Omega|j\rangle]^\dagger$ is just a dual vector whose adjoint is $[\Omega|j\rangle]$ and we know that $\langle v|w\rangle = \langle w|v\rangle^*$. The end result is that one simply transposes and complex conjugates a matrix representation of Ω to get the corresponding matrix representation of $\Omega^\dagger$.

Operator Arithmetic with Adjoints

Because taking the adjoint consists of transposing and complex conjugating, the product rule for adjoint operators is identical to that for transposition:

$$[\Omega\Lambda]^\dagger = \Lambda^\dagger \Omega^\dagger \tag{3.46}$$

which one easily proves by acting on an arbitrary vector:

$$([\Omega\Lambda]\,|v\rangle)^\dagger = (\Omega\,[\Lambda|v\rangle])^\dagger = (\Lambda|v\rangle)^\dagger\,\Omega^\dagger = \left(\langle v\,|\Lambda^\dagger\right)\Omega^\dagger = \langle v\,|\Lambda^\dagger\Omega^\dagger$$

This generalizes our prior rules for dealing with products when taking adjoints of bras and kets: reverse the order of all the factors in a product and take the adjoint of each factor independently.

The adjoint of a sum is of course just the sum of the adjoints.

The above arithmetic rules carry over to the matrix representations of the operators, so taking the adjoint of a product in matrix representation is straightforward. Explicitly, the formula is

$$\left([\Omega\Lambda]^\dagger\right)_{jk} = \left[\Lambda^\dagger\Omega^\dagger\right]_{jk} = \sum_{m=1}^{n}\left[\Lambda^\dagger\right]_{jm}\left[\Omega^\dagger\right]_{mk} = \sum_{m=1}^{n}\Lambda^*_{mj}\Omega^*_{km} = \sum_{m=1}^{n}\Omega^*_{km}\Lambda^*_{mj} \quad (3.47)$$

The next-to-last expression corresponds to: reverse the order of the matrices, complex conjugate and transpose each matrix, then matrix multiply them. The last expression is: complex conjugate each of the two matrices, matrix multiply them, and then transpose the result.

Lecture 7:
Linear Operators Continued
The Eigenvalue-Eigenvector Problem
Date Revised: 2008/10/13
Date Given: 2008/10/13

Hermitian and Anti-Hermitian Operators: the Operators of QM

Hermitian and anti-Hermitian operators are defined by the relations

$$\text{Hermitian: } \Omega^\dagger = \Omega \qquad \text{anti-Hermitian: } \Omega^\dagger = -\Omega \tag{3.48}$$

Remember that, by definition of operator adjoints, the above are equivalent to

$$\begin{aligned} \text{Hermitian:} \quad & \Omega|v\rangle = |w\rangle \Longleftrightarrow \langle v|\Omega = \langle w| \\ \text{anti-Hermitian:} \quad & \Omega|v\rangle = |w\rangle \Longleftrightarrow \langle v|\Omega = -\langle w| \end{aligned} \tag{3.49}$$

The matrix representation versions of the above definitions are

$$\text{Hermitian: } \Omega^*_{jk} = \Omega_{kj} \qquad \text{anti-Hermitian: } \Omega^*_{jk} = -\Omega_{kj} \tag{3.50}$$

A sum of two Hermitian operators is easily seen to be Hermitian. The product of two Hermitian operators *need not* be Hermitian because the two operators have their order reversed when the adjoint is taken: $(\Omega\Lambda)^\dagger = \Lambda^\dagger\Omega^\dagger = \Lambda\Omega \neq \Omega\Lambda$ in general.

Hermitian and anti-Hermitian operators are obvious analogues of purely real and purely imaginary numbers. At a qualitative level, it becomes clear why the operator for a classical physical variable must be Hermitian — we want our physical observables to be real numbers! We will of course justify this rigorously later.

Example 3.28: Projection Operators and Hermiticity

Let us show that our standard projection operator definition is Hermitian. Given a set of orthonormal vectors $\{|j\rangle\}$, $j = 1, \ldots, m$ that span the subspace $\mathbb{V}_P$ of a n-dimensional inner product space that we want to project onto, we have

$$P = \sum_{j=1}^{m} |j\rangle\langle j| \qquad \Longleftrightarrow \qquad P^\dagger = \left(\sum_{j=1}^{m} |j\rangle\langle j| \right)^\dagger = \sum_{j=1}^{m} (|j\rangle\langle j|)^\dagger$$

What is $(|j\rangle\langle j|)^\dagger$? We have an expectation that $(|j\rangle\langle j|)^\dagger = |j\rangle\langle j|$ based on our definition of adjoint for kets and bras, but we have not explicitly showed that should be true when the combination of kets and bras is an operator, not a ket or a bra. So, let's go back to the definition of an operator adjoint, Equation 3.44. In this case, it requires

$$\text{if} \quad (|j\rangle\langle j|)\,|v\rangle = |w\rangle \quad \text{then} \quad \langle w| = \langle v|\,(|j\rangle\langle j|)^\dagger$$

Let's just rewrite the first expression:

$$|j\rangle\langle j|v\rangle = |w\rangle$$

Take the adjoint (which we can do since there are no operators involved now):

$$\langle v\,|j\rangle\langle j\,| = \langle w\,|$$

This looks like the second half of our adjoint definition statement. That statement becomes true if

$$(|j\rangle\langle j\,|)^\dagger = |j\rangle\langle j\,|$$

which is what we expected, but now we have proven it explicitly. So, then,

$$P^\dagger = \sum_{j=1}^{m} (|j\rangle\langle j\,|)^\dagger = \sum_{j=1}^{m} |j\rangle\langle j\,| = P$$

Projection operators are Hermitian.

An alternate definition of a projection operator P is that P be Hermitian and that it satisfy the rather unobvious condition

$$P^2 = P \tag{3.51}$$

Let us show now that this definition is equivalent to our definition Equation 3.29.

First, we show that Equation 3.29 implies the above. We have already demonstrated that it implies Hermiticity. Now let us show it implies Equation 3.51. Again, let $\{|j\rangle\}$, $j = 1, \ldots, m$ be a set of orthonormal vectors that span the subspace $\mathbb{V}_P$ of a n-dimensional inner product space $\mathbb{V}$ that we want to project onto. Then

$$P^2 = \sum_{j=1}^{m} |j\rangle\langle j| \sum_{k=1}^{m} |k\rangle\langle k| = \sum_{j,k=1}^{m} |j\rangle\langle j|k\rangle\langle k| = \sum_{j,k=1}^{m} |j\rangle\delta_{jk}\langle k| = \sum_{j=1}^{m} |j\rangle\langle j| = P$$

as desired.

Linear Operators (cont.)

Let us show the converse, that the conditions $P^2 = P$ and that P is Hermitian imply our original definition Equation 3.29. The condition $P^2 = P$ implies that, for any $|v\rangle$,

$$P(P|v\rangle) = P|v\rangle$$

Let $\mathbb{V}_P$ be the set of vectors produced by acting with P on all $|v\rangle$ belonging to $\mathbb{V}$. We can see that this set is a subspace as follows. Suppose $|v'\rangle$ and $|w'\rangle$ belong to $\mathbb{V}_P$. By definition of $\mathbb{V}_P$, there must be (possibly non-unique) vectors $|v\rangle$ and $|w\rangle$ such that $|v'\rangle = P|v\rangle$ and $|w'\rangle = P|w\rangle$. Then the linear combination $\alpha|v'\rangle + \beta|w'\rangle$ satistfies $\alpha|v'\rangle + \beta|w'\rangle = \alpha P|v\rangle + \beta P|w\rangle = P(\alpha|v\rangle + \beta|w\rangle)$, thereby implying that the linear combination belongs to the $\mathbb{V}_P$ also. So $\mathbb{V}_P$ is closed under all the necessary operations, so it is a subspace.

Now, for any element $|v'\rangle$ in the subspace $\mathbb{V}_P$, it holds that $P|v'\rangle = |v'\rangle$, as follows: For any such element $|v'\rangle$, there is at least one vector $|v\rangle$ such that $|v'\rangle = P|v\rangle$. Since we know $P(P|v\rangle) = P|v\rangle$, it therefore holds $P|v'\rangle = |v'\rangle$.

So, we have that $\mathbb{V}_P$ is a subspace and $P|v\rangle = |v\rangle$ for any $|v\rangle$ in $\mathbb{V}_P$. Let $\{|j\rangle\}$ be an orthonormal basis for this subspace, $j = 1, \ldots, m$ where m is the dimension of the subspace. Then it holds that $P|j\rangle = |j\rangle$ for these $|j\rangle$. Therefore, $\langle k|P|j\rangle = \delta_{kj}$ for $j, k = 1, \ldots, m$. This gives us some of the matrix elements of P.

Extend this orthonormal basis to be an orthonormal basis for the full space, $j = 1, \ldots, n$ where n is the dimension of the full space. We know $\langle j \,|k\rangle = \delta_{jk}$. Therefore, for $j = 1, \ldots, m$ and $k = m+1, \ldots, n$, it holds

$$\langle k \,|\,(P|j\rangle) = |k\rangle\,(|j\rangle) = \langle k \,|j\rangle = \delta_{kj} = 0$$

$$\langle j \,|\,(P|k\rangle) = \langle j \,|P|k\rangle = \left(P^\dagger |j\rangle\right)^\dagger |k\rangle = (P|j\rangle)^\dagger \,|k\rangle = (|j\rangle)^\dagger \,|k\rangle = \langle j \,|k\rangle = \delta_{jk} = 0$$

we used the definition of adjoint operators, bras, and kets and the assumed Hermiticity of P. δ_{jk} vanished in both cases because we had $j = 1, \ldots, m$ and $k = m+1, \ldots, n$: j and k are never the same.

The last matrix elements we need are easy. We want to know what $\langle k \,|P|k\rangle$ is for $k = m+1, \ldots, n$. Since $P|k\rangle$ belongs to $\mathbb{V}_P$ while $|k\rangle$ is orthogonal to the orthonormal basis for $\mathbb{V}_P$, this matrix element always vanishes.

Linear Operators (cont.)

To summarize,

$$\langle j\,|P|k\,\rangle = \begin{cases} \delta_{jk} & \text{for} \quad j,k = 1,\ldots,m \\ 0 & \text{otherwise} \end{cases}$$

We may then use the bilinear form to write out the explicit form for the projection operator:

$$P = \sum_{j,k=1}^{n} |j\,\rangle\langle j\,|P|k\,\rangle\langle k\,| = \sum_{j,k=1}^{m} |j\,\rangle\delta_{jk}\langle k\,| = \sum_{j=1}^{m} |j\,\rangle\langle j\,|$$

which is Equation 3.29, our original definition of the projection operator for the subspace spanned by the orthonormal set $\{|j\,\rangle\}$, $j = 1,\ldots,m$.

We note that $P^2 = P$ does not imply P is its own inverse. Projection operators are in general noninvertible. Let $\mathbb{V}_P$ be the subspace of the inner product space $\mathbb{V}$ onto which the projection operator P projects. Consider a vector $|v\,\rangle$ in the subspace $\mathbb{V}_{P\perp}$ that is orthogonal to $\mathbb{V}_P$, meaning that it is orthogonal to all the $\{|j\,\rangle\}$ comprising P. Then $P|v\,\rangle = |0\,\rangle$. But $P|0\,\rangle = |0\,\rangle$ also, so P is not one-to-one, and hence cannot be invertible. The only case for which this argument fails is for $P = I$ because then the subspace $\mathbb{V}_{P\perp}$ has $|0\,\rangle$ as its only element.

Example 3.29: Spin-1/2 Operators are Hermitian

You can see quite easily that S_x, S_y, and S_z are Hermitian operators by simply taking the complex conjugate transpose of the matrix representations we have already given. For example,

$$S_y^\dagger \leftrightarrow = \left(\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \right)^{*T} = \left(\begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} \right)^{T} = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

Linear Operators (cont.)

Unitary Operators: Operators that do Transformations in QM

Unitary operators are defined by the relation

$$U^\dagger = U^{-1} \tag{3.52}$$

By definition of operator adjoints, the above is equivalent to

$$U|v\rangle = |w\rangle \Longleftrightarrow \langle v|U^{-1} = \langle w| \tag{3.53}$$

We will obtain a definition in terms of matrix representations on the next page.

A product of unitary operators is unitary; one can see this by simply using the product rules for adjoints and inverses. The appropriately normalized sum of unitary operators need not be unitary: when one tests whether $(U_1 + U_2)(U_1 + U_2)^\dagger/4 = I$, one ends up with two cross terms that do not give I unless $U_1 = U_2$.

Unitary operators are like complex numbers of unit modulus, $e^{i\theta}$. Conjugating such a number gives its multiplicative inverse, just as taking the adjoint of a unitary operator gives its operator product inverse. In QM, unitary operators "transform" states — they time evolve them, spatially rotate them, etc. You can think of them as the analogue of the $e^{i\omega t}$ and e^{ikx} factors in electromagnetic wave propagation, though of course their effect is more complicated than that. They are of complementary importance to Hermitian operators.

Inner Products and Unitary Operators

Unitary operators preserve inner products; *i.e.*,

$$\text{if } |v'\rangle = U|v\rangle \text{ and } |w'\rangle = U|w\rangle \text{ then } \langle w'|v'\rangle = \langle w|v\rangle \tag{3.54}$$

The proof is trivial:

$$\langle w'|v'\rangle = (U|w\rangle)^\dagger (U|v\rangle) = \langle w|U^\dagger U|v\rangle = \langle w|v\rangle \tag{3.55}$$

We thus see that unitary operators are generalizations of rotation and other **orthogonal** operators from classical mechanics, which preserve the $\mathbb{R}^3$ dot product.

One property of orthogonal matrices that carries over to unitary operators is the orthonormality of their rows and columns in matrix representation, treating their columns as kets or rows as bras. Shankar gives two proofs; we give the matrix-arithmetic version to provide experience with such manipulations:

$$\langle \text{column } j|\text{column } k\rangle = \sum_m U^*_{mj} U_{mk} = \sum_m \left[U^\dagger\right]_{jm} U_{mk} = \left[U^\dagger U\right]_{jk} = \delta_{jk} \tag{3.56}$$

The row version is similar. Orthonormality of the columns and rows implies the operator is unitary.

Unitary Transformations of Operators

As we noted, unitary operators transform states, such as for time evolution or spatial translation. One of the most basic questions we ask in QM is: how do the matrix elements of some operator change under such a transformation. The interest in the time evolution case is obvious; in other transformations, we are usually interested in how the transformation of operator matrix elements is related to symmetries of the problem.

Explicitly, we might ask: how is $\langle w|\Omega|v\rangle$ related to $\langle w'|\Omega|v'\rangle$ where $|v'\rangle = U|v\rangle$ and $|w'\rangle = U|w\rangle$? Of course the specific answer depends on the problem. But it is generally true that the second expression may be written

$$\langle w'|\Omega|v'\rangle = (U|w\rangle)^\dagger \Omega (U|v\rangle) = \langle w| \left(U^\dagger \Omega U\right) |v\rangle \tag{3.57}$$

The states are now untransformed; instead, we consider the matrix elements of the transformed operator, $\Omega' = U^\dagger \Omega U$ between the untransformed states.

Linear Operators (cont.)

This concept has numerous applications. As we shall see next, we frequently would like to use a basis of eigenstates of some operator H (states $|v\rangle$ for which $H|v\rangle = h|v\rangle$ where h is a number). We can apply a unitary transformation to get from our initial basis to such a basis, and the above transformation lets us see how other operators are represented in the new basis.

Another application is time evolution. The standard picture is the **Schrödinger** picture, in which we apply a unitary time evolution operator to the states. In the alternate **Heisenberg** picture, we leave the states unchanged and apply the time-evolution transformation to operators.

Example 3.30: Particular Unitary Transformations of Spin-1/2 Vectors and Operators

Let's consider a particular set of unitary operators our standard $\mathbb{C}^2$ matrix representation of this space:

$$U = e^{i\,\alpha} \begin{bmatrix} \cos\frac{\theta}{2} & -e^{-i\phi}\sin\frac{\theta}{2} \\ e^{i\,\phi}\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix} \tag{3.58}$$

The four column-row orthonormality conditions leave four degrees of freedom for the arbitary unitary matrix in this representation. This can be represented as one free angle θ that is the argument of the cosines and sines combined with three free phase angles. We have taken one of the phase angles to vanish. Let's try it out on our various states:

$$U(\alpha,\theta,\phi)|\uparrow_z\rangle \leftrightarrow e^{i\,\alpha} \begin{bmatrix} \cos\frac{\theta}{2} & -e^{-i\phi}\sin\frac{\theta}{2} \\ e^{i\,\phi}\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = e^{i\,\alpha} \begin{bmatrix} \cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} \end{bmatrix}$$

$$U(\alpha,\theta,\phi)|\downarrow_z\rangle \leftrightarrow e^{i\,\alpha} \begin{bmatrix} \cos\frac{\theta}{2} & -e^{-i\phi}\sin\frac{\theta}{2} \\ e^{i\,\phi}\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = e^{i\,\alpha} \begin{bmatrix} e^{-i\phi}\sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} \end{bmatrix}$$

In particular, we see that if we take $\theta = \pi$, then we obtain

$$U(\alpha, \theta = \pi, \phi)|\uparrow_z\rangle \leftrightarrow e^{i(\alpha+\phi)} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \leftrightarrow e^{i(\alpha+\phi)}|\downarrow_z\rangle$$

$$U(\alpha, \theta = \pi, \phi)|\downarrow_z\rangle \leftrightarrow e^{i(\alpha-\phi)} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \leftrightarrow e^{i(\alpha-\phi)}|\uparrow_z\rangle$$

This particular unitary operator has rotated the $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$ basis elements into each other, up to unity modulus complex factors. With $\alpha = 0$ and $\phi = 0$, the exchange would be exact. This is equivalent to a spatial rotation of the physical space coordinate axes of π radians about any vector in the xy plane.

Linear Operators (cont.)

What about the action on the other possible bases?

$$U(\alpha=0,\theta=\pi,\phi=0)|\uparrow_x\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}-1\\1\end{bmatrix} \leftrightarrow -|\downarrow_x\rangle$$

$$U(\alpha=0,\theta=\pi,\phi=0)|\downarrow_x\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix} \leftrightarrow |\uparrow_x\rangle$$

$$U(\alpha=0,\theta=\pi,\phi=0)|\uparrow_y\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix}1\\i\end{bmatrix} = -i\begin{bmatrix}1\\i\end{bmatrix} \leftrightarrow -i\,|\uparrow_y\rangle$$

$$U(\alpha=0,\theta=\pi,\phi=0)|\downarrow_y\rangle \leftrightarrow \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-i\end{bmatrix} = i\begin{bmatrix}1\\-i\end{bmatrix} \leftrightarrow i\,|\downarrow_y\rangle$$

We see that the physical space rotation is about the y axis, so that the transformation rotates $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$ into each other (modulo signs) and keeps $|\uparrow_y\rangle$ and $|\downarrow_y\rangle$ unchanged (modulo unity modulus factors).

Linear Operators (cont.)

How about unitary transformation of operators? Again, using $\alpha = 0$, $\theta = \pi$, and $\phi = 0$, let's apply the unitary transformation to S_x, S_y, and S_z:

$$U^\dagger S_z U \leftrightarrow \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \frac{\hbar}{2} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \leftrightarrow -S_z$$

$$U^\dagger S_x U \leftrightarrow \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \frac{\hbar}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \frac{\hbar}{2} \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \leftrightarrow -S_x$$

$$U^\dagger S_y U \leftrightarrow \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \frac{\hbar}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \frac{\hbar}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \leftrightarrow S_y$$

The sign flips on S_z and S_x make sense, as we saw that the corresponding basis states were rotated into each other, while the lack of change for S_y makes sense because we saw the unitary transformation left them unaffected except for prefactors of unity modulus.

Note that one *either* transforms the states *or* transforms the operators, not both — that's why the exchange of $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$ and the sign flip on S_z do not cancel one another because one does not both of them, one does only one, depending on whether you want to transform the states or the operators.

The Eigenvector-Eigenvalue Problem: Formalism

Motivation

An eigenvector of an operator is a vector that is left unchanged by the operator up to a scalar multiplier, which is called the eigenvalue. The vector $|\omega\rangle$ is an eigenvector of the operator Ω with eigenvalue ω if and only if

$$\Omega|\omega\rangle = \omega|\omega\rangle \tag{3.59}$$

The key point is that the eigenvector's direction in the inner product space is left unchanged by the action of the operator. An operator can have multiple eigenvectors, and the eigenvectors need not all be different.

One of the postulates of QM is that measurement of any classical variable yields only the eigenvalues of the corresponding quantum operator, with only the probability of obtaining any particular value known ahead of time, and that the act of measuring the physical quantity results in collapse of the state to the eigenstate corresponding to the measured eigenvalue.

It is therefore not surprising that we must study the problem of eigenvalues and eigenvalues in inner product spaces.

You have seen material of this type repeatedly, in your discussion of both normal modes and of quantum mechanics in Ph2/12. As usual, though, we will proceed methodically to ensure you understand the eigenvalue-eigenvector problem deeply.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

Statement of the Problem

Given a linear operator Ω. How do we find its eigenvalues and eigenvectors? We are asking for solutions to the linear equation (I is the identity operator)

$$\Omega|v\rangle = \omega|v\rangle \quad \Longleftrightarrow \quad (\Omega - I\,\omega)\,|v\rangle = |0\rangle \tag{3.60}$$

Solution of the Problem

You know from studying linear algebra that the above equation is only true if the determinant of any matrix representation of the operator on the left side vanishes:

$$|\Omega - I\,\omega| = 0 \tag{3.61}$$

This equation is termed the characteristic equation for Ω.

Here we begin to get sloppy about the difference between equality $=$ and representation $\leftrightarrow$: a determinant only makes sense for a matrix representation, not for an operator, but we are using the symbol for the operator in the above equation. We could dream up some notation to distinguish between the operator and a matrix representation of it; for example, Ω and $[\Omega]$ or Ω and $\underline{\Omega}$. This will become very tedious to carry along for the remainder of the course, though, so from here on we will have to rely on context to distinguish between an operator and its matrix representation.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

To properly justify Equation 3.61, one must: 1) prove a general formula for the inverse of a matrix when a matrix representation is specified; 2) assume that the operator $\Omega - I\,\omega$ is noninvertible so that $[\Omega - I\,\omega]\,|v\rangle = |0\rangle$ does not imply $|v\rangle = |0\rangle$; and 3) use noninvertiblity and the inversion formula to obtain $|\Omega - I\,\omega| = 0$. See Shankar Appendix A.1, Equation A.1.7 and Theorem A.1.1.

The formula only can be written explicitly when a matrix representation is specified for Ω and $|v\rangle$, which is only possible when an orthonormal basis is specified. Let's assume this has been given. Then we can write out the determinant. Since we have put no conditions on Ω, all we can say at this point is that the the resulting equation is a nth-order polynomial in ω where n is the dimension of the space: the diagonal of $\Omega - I\,\omega$ has one power of ω in each element, and the determinant will include one term that is the product of all these elements, so there is at least one term in ω^n. So, the eigenvalues will be given by the solution to the polynomial equation

$$p_n(\omega) = \sum_{m=0}^{n} c_m \omega^m = 0 \tag{3.62}$$

The polynomial p_n is called the characteristic polynomial for the operator Ω. The fundamental theorem of algebra tells us it has n roots, some possibly complex. If the vector space's field is complex, then these are valid eigenvalues; if the field were real, then we say that some of the roots do not exist. Thus, any linear operator in a vector space whose field is the complex numbers is guaranteed to have as many eigenvalues as the dimension of the vector space. Since the eigenvalues are independent of the basis and the matrix representation (Equation 3.60 is basis- and representation-independent), the characteristic polynomial must also be.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

Once we have the eigenvalues, how do we find the eigenvectors?

Easy: for a particular eigenvalue ω_j and eigenvector $|\omega_j\rangle$, we have the equation

$$(\Omega - I\,\omega_j)\,|\omega_j\rangle = |0\rangle \tag{3.63}$$

Since we explicitly know what the operator is — we know the elements of Ω and we know ω_j — all we need to do is solve for the elements of $|\omega_j\rangle$. Formally, though, because the determinant of the matrix on the left vanishes, we are not guaranteed a unique solution. What we end up with is $n-1$ independent linear equations that determine $n-1$ components of $|\omega_j\rangle$, leaving the overall normalization undetermined. The normalization of $|\omega_j\rangle$ is arbitrary since, if $|\omega_j\rangle$ is an eigenvector, then $\alpha|\omega_j\rangle$ will also be an eigenvector for any α.

Of course, if our vector space has a real field, which may result in some of the eigenvalues not existing in the field, then the corresponding eigenvectors will also not exist because we would simply not be allowed to write Equation 3.63 for that eigenvalue.

In some cases, the above procedure will not yield the n eigenvectors in that one will obtain $|\omega_j\rangle = |0\rangle$; this happens when there are degenerate (equal) eigenvalues. We can prove some explicit theorems about the existence of eigenvectors and the nature of the eigenvalues when the operators are Hermitian or unitary, which we will do below.

Example 3.31: In $\mathbb{C}^3$, rotation about the vector $(\widehat{x}+\widehat{y}+\widehat{z})/\sqrt{3}$, which simply cyclically permutes the three unit vectors.

The matrix representation of this operator, which rotates $\widehat{x} \longrightarrow \widehat{y}$, $\widehat{y} \longrightarrow \widehat{z}$ and $\widehat{z} \longrightarrow \widehat{x}$, is

$$A \leftrightarrow \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

The characteristic equation is

$$0 = |A - I\,\omega| = \begin{vmatrix} -\omega & 0 & 1 \\ 1 & -\omega & 0 \\ 0 & 1 & -\omega \end{vmatrix} = -\omega^3 + 1 = 0$$

$$\omega_1 = 1 \qquad \omega_2 = e^{2\pi i/3} \qquad \omega_3 = e^{-2\pi i/3}$$

If we had assumed $\mathbb{R}^3$, we would say that two of the eigenvalues and the corresponding eigenvectors do not exist.

Let us find the eigenvectors for each case by calculating $A - I\,\omega$ for each case and solving $(A - I\,\omega)\,|v\rangle = |0\rangle$.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

$\omega_1 = 1$:

$$\begin{bmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{12} \\ v_{13} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \implies \begin{aligned} -v_{11} + v_{13} &= 0 \\ v_{11} - v_{12} &= 0 \\ v_{12} - v_{13} &= 0 \end{aligned}$$

$$\implies |\omega_1\rangle \leftrightarrow \alpha \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

As expected, the normalization is not set because the three equations are not independent. The conventional choice is to normalize to 1, so in this case $\alpha = 1/\sqrt{3}$. As one would expect, the vector corresponding to the axis of rotation has eigenvalue 1.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

$\omega_2 = e^{\frac{2\pi i}{3}}$:

$$\begin{bmatrix} -e^{\frac{2\pi i}{3}} & 0 & 1 \\ 1 & -e^{\frac{2\pi i}{3}} & 0 \\ 0 & 1 & -e^{\frac{2\pi i}{3}} \end{bmatrix} \begin{bmatrix} v_{21} \\ v_{22} \\ v_{23} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Longrightarrow \begin{array}{r} -e^{\frac{2\pi i}{3}} v_{21} + v_{23} = 0 \\ v_{21} - e^{\frac{2\pi i}{3}} v_{22} = 0 \\ v_{22} - e^{\frac{2\pi i}{3}} v_{23} = 0 \end{array}$$

$$\Longrightarrow \quad |\omega_2\rangle \leftrightarrow \alpha \begin{bmatrix} 1 \\ e^{-\frac{2\pi i}{3}} \\ e^{\frac{2\pi i}{3}} \end{bmatrix}$$

Since all the elements are unit modulus, we again may take $\alpha = 1/\sqrt{3}$. Had we restricted to $\mathbb{R}^3$, we would have said this eigenvector did not exist (which makes sense, given that the eigenvalue would not have been in the scalar field of the vector space).

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

$\omega_3 = e^{-\frac{2\pi i}{3}}$:

$$\begin{bmatrix} -e^{-\frac{2\pi i}{3}} & 0 & 1 \\ 1 & -e^{-\frac{2\pi i}{3}} & 0 \\ 0 & 1 & -e^{-\frac{2\pi i}{3}} \end{bmatrix} \begin{bmatrix} v_{31} \\ v_{32} \\ v_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \implies \begin{aligned} -e^{-\frac{2\pi i}{3}} v_{31} + v_{33} &= 0 \\ v_{31} - e^{-\frac{2\pi i}{3}} v_{32} &= 0 \\ v_{32} - e^{-\frac{2\pi i}{3}} v_{33} &= 0 \end{aligned}$$

$$\implies |\omega_3\rangle \leftrightarrow \alpha \begin{bmatrix} 1 \\ e^{\frac{2\pi i}{3}} \\ e^{-\frac{2\pi i}{3}} \end{bmatrix}$$

Again, we may take $\alpha = 1/\sqrt{3}$, and, again, this eigenvector would be said to not exist if we had restricted to $\mathbb{R}^3$.

Lecture 8:
The Eigenvector-Eigenvalue Problem Continued

Date Revised: 2008/10/15
Date Given: 2008/10/15

Degeneracy

What happens when two or more eigenvalues are equal? Intuitively, one sees that, if there were two eigenvectors $|\omega, 1\rangle$ and $|\omega, 2\rangle$ corresponding to the same eigenvalue ω, then any linear combination would also be an eigenvector with the same eigenvalue:

$$\text{if} \quad A|\omega,1\rangle = \omega|\omega,1\rangle \quad \text{and} \quad A|\omega,2\rangle = \omega|\omega,2\rangle \tag{3.64}$$
$$\text{then} \quad A\left(\alpha|\omega,1\rangle + \beta|\omega,2\rangle\right) = \alpha\,\omega|\omega,1\rangle + \beta\,\omega|\omega,2\rangle = \omega\left(\alpha|\omega,1\rangle + \beta|\omega,2\rangle\right)$$

Hence, one expects that the formalism should be unable to pick between $|\omega, 1\rangle$, $|\omega, 2\rangle$, and any linear combination of the two. It in fact does have problems; in general, rather than there being just one redundant equation when one solves for the eigenvector, there are n_d redundant equations where n_d is the number of degenerate eigenvalues. This is to be expected, as what the problem is saying is that all vectors in a subspace of dimension n_d are eigenvectors, and it's therefore entirely arbitrary which n_d of those vectors one chooses to be the nominal eigenvectors. Of course, if one wants to span the subspace, one had better pick linearly independent ones.

We will show below that any pair of eigenvectors corresponding to nondegenerate eigenvalues are always orthogonal. Motivated by this, the usual procedure is to pick a convenient set of orthogonal vectors in the degenerate subspace as the eigenvectors. They are automatically orthogonal to the other, nondegenerate eigenvectors, and making them orthogonal provides an overall orthogonal (and hence easily orthonormalizable) basis for the inner product space.

Theorems on Properties of Eigenvalues and Eigenvectors

The eigenvalues of a Hermitian operator are real.

Assume the Hermitian operator Ω has eigenvalue ω with eigenvector $|\omega\rangle$, $\Omega|\omega\rangle = \omega|\omega\rangle$. Take the matrix element of Ω between the ket $|\omega\rangle$ and bra $\langle\omega|$ (also known as the expectation value of Ω as we shall see later):

$$\langle\omega|\Omega|\omega\rangle = \omega\langle\omega|\omega\rangle \tag{3.65}$$

Also consider the adjoint of the above expression

$$\langle\omega|\Omega^\dagger|\omega\rangle = \omega^*\langle\omega|\omega\rangle \tag{3.66}$$

The two expressions must be equal because $\Omega^\dagger = \Omega$, so we have

$$(\omega - \omega^*)\langle\omega|\omega\rangle = 0 \tag{3.67}$$

Unless $\langle\omega|\omega\rangle = 0$, which can only hold for $|\omega\rangle = |0\rangle$, implying a trivial operator Ω, we find that $\omega = \omega^*$; *i.e.*, the eigenvalue ω is real.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

Any pair of eigenvectors corresponding to nondegenerate eigenvalues of a Hermitian operator are orthogonal.

Given two eigenvalues ω_j and ω_k and corresponding eigenvectors $|\omega_j\rangle$ and $|\omega_k\rangle$, we have

$$\langle\omega_j|\Omega|\omega_k\rangle = \langle\omega_j|\omega_k|\omega_k\rangle = \omega_k\langle\omega_j|\omega_k\rangle \tag{3.68}$$

and

$$\begin{aligned}\langle\omega_j|\Omega|\omega_k\rangle &= \left(\Omega^\dagger|\omega_j\rangle\right)^\dagger|\omega_k\rangle = \left(\Omega|\omega_j\rangle\right)^\dagger|\omega_k\rangle = \left(\omega_j|\omega_j\rangle\right)^\dagger|\omega_k\rangle = \langle\omega_j|\omega_j^*|\omega_k\rangle \\ &= \langle\omega_j|\omega_j|\omega_k\rangle = \omega_j\langle\omega_j|\omega_k\rangle\end{aligned} \tag{3.69}$$

where we have used that Ω is Hermitian and that its eigenvalue ω_j is real. We thus have

$$\left(\omega_j - \omega_k\right)\langle\omega_j|\omega_k\rangle = 0 \tag{3.70}$$

Because we assumed nondegenerate eigenvalues $\omega_j \neq \omega_k$, we have $\langle\omega_j|\omega_k\rangle = 0$.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

For any Hermitian operator acting on an inner product space with a complex field, there exists an orthonormal basis of its eigenvectors, termed its eigenbasis.

We will first prove this for the case of no degenerate eigenvalues. Our proof is somewhat different than Shankar's.

The proof is almost trivial. Any Hermitian operator acting on a n-dimensional inner product space with a complex field has n eigenvalues because the operator has a $n \times n$ matrix representation, yielding a characteristic polynomial of nth order. As mentioned before, it is guaranteed to have n complex roots. There are thus n eigenvalues, nondegenerate by assumption here.

We have shown that, for nondegenerate eigenvalues, the eigenvectors of any pair of eigenvalues are orthogonal. We are thus assured of a mutually orthogonal set of n eigenvectors. It is trivial to render these orthonormal by picking their normalization appropriately (the length of an eigenvector is arbitrary, recall).

Finally, because our orthonormal set is clearly linearly independent, and because it contains n vectors, it is a valid basis for the n-dimensional inner product space.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

When represented in terms of a basis of its eigenvectors, a Hermitian operator's matrix representation is diagonal and its diagonal elements are its eigenvalues.

Again, we take a different tack than Shankar. If we write Ω in a matrix representation in which its eigenvectors are the basis, then its eigenvectors have matrix representation

$$|\omega_1\rangle \leftrightarrow \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad |\omega_2\rangle \leftrightarrow \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \cdots \quad |\omega_n\rangle \leftrightarrow \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

The matrix representation of an operator in a particular basis's matrix representation is given by the matrix elements of the operator between the basis members according to Equation 3.37. So, here we have

$$\Omega \leftrightarrow \begin{bmatrix} \langle\omega_1|\Omega|\omega_1\rangle & \cdots & \langle\omega_1|\Omega|\omega_n\rangle \\ \vdots & \ddots & \vdots \\ \langle\omega_n|\Omega|\omega_1\rangle & \cdots & \langle\omega_n|\Omega|\omega_n\rangle \end{bmatrix} = \begin{bmatrix} \omega_1 & 0 & \cdots & 0 \\ 0 & \omega_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \omega_n \end{bmatrix} \tag{3.71}$$

because the basis elements are eigenvectors of Ω and form an orthonormal set; that is, because $\Omega_{jk} = \langle\omega_j|\Omega|\omega_k\rangle = \omega_k\langle\omega_j|\omega_k\rangle = \omega_k\delta_{jk}$.

Recalling our bilinear form for operators, Equation 3.42, we may also use the condition $\Omega_{jk} = \omega_j \delta_{jk}$ on the matrix elements of Ω in its eigenbasis to write the operator in the form

$$\Omega = \sum_{j,k=1}^{n} |\omega_j\rangle\langle\omega_j|\Omega|\omega\rangle k\langle\omega_k| = \sum_{j,k=1}^{n} |\omega_j\rangle\omega_j\delta_{jk}\langle\omega_k| = \sum_{j=1}^{n} \omega_j|\omega_j\rangle\langle\omega_j| \qquad (3.72)$$

This makes it explicit that Ω's matrix representation is diagonal when the basis for the matrix representation is Ω's eigenbasis.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

Degenerate case:

Even if one has degenerate eigenvalues, the above results still hold – one can still construct an orthonormal basis of the operator's eigenvectors, and then one can write the matrix representation of the operator and it is diagonal. We are not going to be strictly rigorous about proving this, but we can make a fairly ironclad argument.

Let ω be an eigenvalue that is n_d times degenerate. We know that the set of vectors that are eigenvectors with this eigenvalue form a subspace because the set is closed under linear combinations, as we noted earlier (the other arithmetic properties of the subspace are inherited from the parent space.)

Let us assume for the moment that ω is the only degenerate eigenvalue, so that there are $n_n = n - n_d$ nondegenerate eigenvalues. This provides n_n mutually orthogonal eigenvectors as shown above. Note also that our eigenvector orthogonality proof also implies that these nondegenerate eigenvectors are orthogonal to any vector in the ω subspace because any vector in that subspace is an eigenvector of Ω with eigenvalue ω, which is a different eigenvalue from any of the n_n nondegenerate eigenvalues, and hence the previously given proof of orthogonality carries through.

We thus have a n-dimensional vector space with a subspace of dimension $n_n = n - n_d$. We make the intuitively obvious claim that the remaining subspace, which is the degenerate subspace, thus has dimension n_d and therefore has at least one linearly independent basis set with n_d elements.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

Finally, we invoke Gram-Schmidt orthogonalization to turn that linearly independent basis into an orthonormal basis. This basis for the degenerate subspace is automatically orthogonal to the eigenvectors with nondegenerate eigenvalues, so together they form an orthonormal basis for the entire space.

If there is more than one degenerate eigenvalue, one simply performs the above procedure for each degenerate subspace independently.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

The eigenvalues of a unitary operator are complex numbers of unit modulus.

Consider the norm of an eigenvector $|\omega\rangle$ of the unitary operator with eigenvalue ω:

$$\langle\omega|\omega\rangle = \langle\omega|U^\dagger U|\omega\rangle = \langle\omega|\omega^*\omega|\omega\rangle \Longrightarrow (\omega^*\omega - 1)\langle\omega|\omega\rangle = 0 \tag{3.73}$$

For nontrivial $|\omega\rangle$, we have $\omega^*\omega = 1$, and hence ω must have unit modulus. (This last step you can prove by writing ω out in terms of real and imaginary components and solving the equation for its components.)

The eigenvectors of a unitary operator are mutually orthogonal.

Consider a similar construct, this time the inner product of eigenvectors $|\omega_j\rangle$ and $|\omega_k\rangle$ of two nondegenerate eigenvalues $\omega_j \neq \omega_k$:

$$\langle\omega_j|\omega_k\rangle = \langle\omega_j|U^\dagger U|\omega_k\rangle = \langle\omega|\omega_j^*\omega_k|\omega\rangle \Longrightarrow \left(\omega_j^*\omega_k - 1\right)\langle\omega_j|\omega_k\rangle = 0 \tag{3.74}$$

For $\omega_j \neq \omega_k$, the quantity $\omega_j^*\omega_k - 1$ cannot vanish unless $\omega_j = \omega_k$, which we assumed did not hold. Therefore $\langle\omega_j|\omega_k\rangle = 0$ and we have orthogonality.

Of course, we can deal with degenerate subspaces in the same way as we did for Hermitian operators.

Diagonalization of Hermitian Matrices and Unitary Transformations

Since we have shown that one can always construct an orthonormal basis of the eigenvectors of a Hermitian matrix, we can write down a unitary operator whose matrix representation **in the original basis** $\{|j\rangle\}$ is made up from the components of those eigenvectors in that basis:

$$U_\Omega \leftrightarrow \begin{bmatrix} \langle 1|\omega_1\rangle & \cdots & \langle 1|\omega_n\rangle \\ \vdots & \ddots & \vdots \\ \langle n|\omega_1\rangle & \cdots & \langle n|\omega_n\rangle \end{bmatrix} \tag{3.75}$$

That is, we make up each column of the matrix from the expansion coefficients of the eigenvectors in the original basis $\{|j\rangle\}$: the first column contains the expansion coefficients of $|\omega_1\rangle$ in that basis, the second column contains those of $|\omega_2\rangle$ and so on.

Let's check that it satisfies the column-wise and row-wise orthonormality conditions required of a unitary operator's matrix representation. First, the column-wise proof:

$$\langle \text{column } j \,|\text{column } k \,\rangle = \sum_{m=1}^{n} \langle m \,|\omega_j \,\rangle^* \langle m \,|\omega_k \,\rangle = \sum_{m=1}^{n} \langle \omega_j \,|m \,\rangle \langle m \,|\omega_k \,\rangle = \langle \omega_j \,|\omega_k \,\rangle = \delta_{jk}$$

where we used the fact that the $\{|j\,\rangle\}$ are an orthonormal basis that span the space and that the $\{|\omega_j\,\rangle\}$ are an orthonormal set. Similarly, for the row-wise condition:

$$\langle \text{row } j \,|\text{row } k \,\rangle = \sum_{m=1}^{n} \langle j \,|\omega_m \,\rangle^* \langle k \,|\omega_m \,\rangle = \sum_{m=1}^{n} \langle k \,|\omega_m \,\rangle \langle \omega_m \,|j \,\rangle = \langle k \,|j \,\rangle = \delta_{kj}$$

where now we use the fact that the $\{|\omega_j\,\rangle\}$ are an orthonormal basis that span the space and that the $\{|j\,\rangle\}$ are an orthonormal set.

What does this unitary operator do? If we act on the column matrix representation of one of the basis elements that defines the matrix representation, we find it gets transformed into one of the eigenvectors; for example, acting on $|n\rangle$:

$$U_\Omega|n\rangle \leftrightarrow \begin{bmatrix} \langle 1|\omega_1\rangle & \cdots & \langle 1|\omega_n\rangle \\ \vdots & \ddots & \vdots \\ \langle n|\omega_1\rangle & \cdots & \langle n|\omega_n\rangle \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} \langle 1|\omega_n\rangle \\ \vdots \\ \langle n|\omega_n\rangle \end{bmatrix}$$

$$\leftrightarrow \sum_{j=1}^{n} \langle j|\omega_n\rangle |j\rangle = \sum_{j=1}^{n} |j\rangle\langle j|\omega_n\rangle = |\omega_n\rangle$$

where we again used the fact that the $\{|j\rangle\}$ are an orthonormal basis for the space to collapse the sum over j. Similarly, the reverse transformation from the eigenvectors to the original basis is performed by $U_\Omega^\dagger$; we summarize these two statements as

$$|\omega_j\rangle = U_\Omega|j\rangle \qquad |j\rangle = U_\Omega^\dagger|\omega_j\rangle \tag{3.76}$$

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

If U_Ω rotates the original orthonormal basis $\{|j\rangle\}$ to become the eigenvectors $\{|\omega_j\rangle\}$, and $U_\Omega^\dagger$ rotates the eigenvectors $\{|\omega_j\rangle\}$ to become the original orthonormal basis $\{|j\rangle\}$, we are led to the question: how does U_Ω act on the operator Ω that gave the eigenvectors? Consider the following:

$$\langle\omega_j|\Omega|\omega_k\rangle = \langle\omega_j|U_\Omega\, U_\Omega^\dagger\, \Omega\, U_\Omega U_\Omega^\dagger|\omega_k\rangle = \langle j|U_\Omega^\dagger\, \Omega\, U_\Omega|k\rangle \tag{3.77}$$

Since we know $\langle\omega_j|\Omega|\omega_k\rangle = \omega_j\delta_{jk}$, it must therefore hold that the unitary transformation

$$\Omega' = U_\Omega^\dagger\, \Omega\, U_\Omega \tag{3.78}$$

gives a new operator Ω' that is diagonal in the original basis $\{|j\rangle\}$ and has the same eigenvalues, in the same order, as Ω:

$$\langle\omega_j|\Omega|\omega_k\rangle = \langle j|\Omega'|k\rangle \tag{3.79}$$

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

More generally, given an operator Λ and the unitary operator U_Ω, we define the transformed version Λ' in the same manner

$$\Lambda' = U_\Omega^\dagger \, \Lambda \, U_\Omega \tag{3.80}$$

Whether or not this transformation gives an operator Λ' that is diagonal in the original basis $\{|j\rangle\}$ like Ω' depends on Λ; we will return to this question soon.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

We must explain a subtle point about "transformations" versus "arithmetic operations." The above discussion gives us a **new operator** $\Omega' = U_\Omega^\dagger \, \Omega \, U_\Omega$ that is **different from** Ω! The original operator Ω is diagonal if one's matrix representation uses the $\{|\omega_j\rangle\}$ as the basis; the new operator is diagonal if one's matrix representation uses the original $\{|j\rangle\}$ basis. They are different operators and they have different eigenvectors.

One thing that is confusing about this is that the operators Ω and Ω' have the same eigenvalues and thus their diagonal forms are the same. One is tempted to think that they are the same operator. But, because they are diagonal in different matrix representations, they are most definitely not the same operator. An explicit way to see this is to write them out in the form given in Equation 3.72

$$\Omega = \sum_{j=1}^{n} \omega_j |\omega_j\rangle\langle\omega_j| \qquad \Omega' = \sum_{j=1}^{n} \omega_j |j\rangle\langle j|$$

The two forms involve outer products of entirely different sets of vectors, so they are different operators; it is only the coefficients that are the same.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

We can see that each operator is diagonal in its own eigenbasis and not diagonal in the other's eigenbasis by writing out the relevant matrix elements:

$$\langle \omega_j | \Omega | \omega_k \rangle = \langle \omega_j | \left[\sum_{m=1}^{n} \omega_m |\omega_m\rangle\langle\omega_m| \right] |\omega_k\rangle = \omega_j \delta_{jk}$$

$$\langle j | \Omega | k \rangle = \langle j | \left[\sum_{m=1}^{n} \omega_m |\omega_m\rangle\langle\omega_m| \right] |k\rangle = \left[\sum_{m=1}^{n} \omega_m \langle j|\omega_m\rangle\langle\omega_m|k\rangle \right]$$

$$\langle \omega_j | \Omega' | \omega_k \rangle = \langle \omega_j | \left[\sum_{m=1}^{n} \omega_m |m\rangle\langle m| \right] |\omega_k\rangle = \left[\sum_{m=1}^{n} \omega_m \langle \omega_j|m\rangle\langle m|\omega_k\rangle \right]$$

$$\langle j | \Omega' | k \rangle = \langle j | \left[\sum_{m=1}^{n} \omega_m |m\rangle\langle m| \right] |k\rangle = \omega_j \delta_{jk}$$

The matrix elements are diagonal for each operator in its own eigenbasis, but are not necessarily diagonal for each operator in the other operator's eigenbasis. We also see we recover Equation 3.77,

$$\langle \omega_j | \Omega | \omega_k \rangle = \omega_j \delta_{jk} = \langle j | \Omega' | k \rangle$$

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

There is something that muddles all of this. To obtain the representation of an arbitary operator Λ in the matrix representation corresponding to the eigenbasis of Ω, $\{|\omega_j\rangle\}$, we find we must apply the unitary transformation operation $U_\Omega^\dagger \Lambda U$ to the matrix representation of Λ in the original basis $\{|j\rangle\}$ **as a purely arithmetic procedure**. At the cost of some notational complexity, we can clarify the similarity and difference between the unitary transformation as an operator transformation, yielding a new operator Λ', and its use as an arithmetic procedure to obtain a new matrix representation of the same operator Λ. Let us use the following notation:

$$\Lambda = \text{an operator on a vector space, representation-free}$$

$$\left[\ \Lambda\ \right]_{|j\rangle} = \begin{array}{l}\text{the matrix representation of the operator } \Lambda \\ \text{in the } \{|j\rangle\} \text{ matrix representation}\end{array}$$

$$\left[\ \Lambda\ \right]_{|\omega_j\rangle} = \begin{array}{l}\text{the matrix representation of the operator } \Lambda \\ \text{in the } \{|\omega_j\rangle\} \text{ matrix representation}\end{array}$$

Next, note the following relationship between matrix elements in the two different bases:

$$\langle \omega_j \,|\Lambda|\omega_k \,\rangle = \sum_{p,q=1}^{n} \langle \omega_j \,|p\,\rangle\langle p\,|\Lambda|q\,\rangle\langle q\,|\omega_k\,\rangle = \sum_{p,q=1}^{n} \left[U_\Omega^\dagger\right]_{jp} \langle p\,|\Lambda|q\,\rangle\,[U_\Omega]_{qk}$$

So, **the matrix elements are related by an arithmetic operation that is the same as the unitary transformation**. Using our notation,

$$\left[\ \Lambda\ \right]_{|\omega_j\,\rangle} = \left[\ U_\Omega^\dagger\ \right]_{|j\,\rangle} \left[\ \Lambda\ \right]_{|j\,\rangle} \left[\ U_\Omega\ \right]_{|j\,\rangle} \tag{3.81}$$

But we also have, based on $\Lambda' = U_\Omega^\dagger \Lambda U_\Omega$:

$$\left[\ \Lambda'\ \right]_{|j\,\rangle} = \left[\ U_\Omega^\dagger\ \right]_{|j\,\rangle} \left[\ \Lambda\ \right]_{|j\,\rangle} \left[\ U_\Omega\ \right]_{|j\,\rangle} \tag{3.82}$$

We may therefore state

$$\left[\ \Lambda'\ \right]_{|j\rangle} = \left[\ \Lambda\ \right]_{|\omega_j\rangle} \tag{3.83}$$

which is no doubt incredibly confusing: the matrix representation of the unitary-transformed operator Λ' in the original $\{|j\rangle\}$ basis is the same as the matrix representation of the untransformed operator Λ in the eigenbasis $\{|\omega_j\rangle\}$. Thus, one has to be very careful to understand from context whether one is staying in the same basis and transforming the operators or whether one is going to the matrix representation of the eigenbasis. The matrix representations will look the same! We usually want to do the latter, but in practice do the former.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

Going back to Equation 3.77, let's consider another possible source of confusion. Is it completely clear what we mean by U_Ω and $\left[\ U_\Omega\ \right]_{|j\rangle}$? We defined U_Ω by its matrix representation in the $\{|j\rangle\}$ basis, and that is what we mean above by $\left[\ U_\Omega\ \right]_{|j\rangle}$ and its adjoint. That's clear and unambiguous.

But confusion may arise when one asks the obvious follow-on question to: what should the matrix representation of U_Ω in Ω's eigenbasis, $\left[\ U_\Omega\ \right]_{|\omega_j\rangle}$, be? Should it be the identity matrix because no transformation is needed if one's matrix representation is already in the eigenbasis of Ω? Applying Equation 3.77, we obtain

$$\left[\ U_\Omega\ \right]_{|\omega_j\rangle} = \left[\ U_\Omega^\dagger\ \right]_{|j\rangle} \left[\ U_\Omega\ \right]_{|j\rangle} \left[\ U_\Omega\ \right]_{|j\rangle} = \left[\ U_\Omega\ \right]_{|j\rangle}$$

which indicates that U_Ω has the same matrix representation in the two bases. Which is correct: is U_Ω's matrix representation independent of basis, or should U_Ω become the identity matrix in the eigenbasis of Ω?

The confusion arises because we have been a bit ambiguous about what is meant by U_Ω. U_Ω is the operator that transforms $\{|j\rangle\}$ into $\{|\omega_j\rangle\}$. This depends on both $\{|j\rangle\}$ and $\{|\omega_j\rangle\}$, not just on $\{|\omega_j\rangle\}$ (and thus not just on Ω). Really, we ought to label U_Ω as $U_{|j\rangle\rightarrow|\omega_j\rangle}$ because U is defined in terms of $\{|j\rangle\}$ and $\{|\omega_j\rangle\}$; it depends only indirectly on Ω through the fact that Ω determines what the $\{|\omega_j\rangle\}$ are. If one's basis is already the eigenbasis of Ω, then the unitary operator one wants is $U_{|\omega_j\rangle\rightarrow|\omega_j\rangle} = I$. That is a different operator from $U_{|j\rangle\rightarrow|\omega_j\rangle}$.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

Thus, strictly speaking, the preceding equation that relates the matrix elements of U_Ω in the two bases should be written as

$$\begin{aligned}\left[\ U_{|j\rangle\to|\omega_j\rangle}\ \right]_{|\omega_j\rangle} &= \left[\ U^\dagger_{|j\rangle\to|\omega_j\rangle}\ \right]_{|j\rangle}\left[\ U_{|j\rangle\to|\omega_j\rangle}\ \right]_{|j\rangle}\left[\ U_{|j\rangle\to|\omega_j\rangle}\ \right]_{|j\rangle} \\ &= \left[\ U_{|j\rangle\to|\omega_j\rangle}\ \right]_{|j\rangle} \end{aligned} \tag{3.84}$$

The matrix representation of $U_{|j\rangle\to|\omega_j\rangle}$ is indeed unchanged by the unitary transformation. The resolution of the misconception is that this is no longer the operator one wants: if working in the $|\omega_j\rangle$ basis, one wants $U_{|\omega_j\rangle\to|\omega_j\rangle} = I$. The above form for $\left[\ U_{|j\rangle\to|\omega_j\rangle}\ \right]_{|\omega_j\rangle}$ is therefore not wrong, it is simply not useful. It now would rotate the eigenbasis of Ω to some new set of vectors in the space that are neither the original basis nor the eigenbasis of Ω.

Clearly, there is much opportunity for confusion. We cannot use the above notation in general because it is too complicated to carry around. We will have to rely on context to understand which U_Ω we are interested in. One saving grace, though, is that, once we have decided which bases U_Ω will transform between, then its matrix representation is independent of basis choice. Therefore, we will not need to write $\left[\ U_\Omega\ \right]_{|j\rangle}$, we can simply write U_Ω.

Lecture 9:
The Eigenvector-Eigenvalue Problem Continued
Unitary Transformations Revisited
Functions of Operators
Calculus with Operators

Date Revised: 2008/10/17
Date Given: 2008/10/17 and 2008/10/20

The Eigenvector-Eigenvalue Problem: Formalism

If Ω and Λ are two commuting Hermitian operators, there is a basis of common eigenvectors in whose matrix representation both operators are diagonal.

As usual, first let's do the nondegenerate case. Assume that Ω has no degenerate eigenvalues, and that its eigenvalues and eigenvectors are, as usual, $\{\omega_j\}$ and $\{|\omega_j\rangle\}$. Then we have

$$\Lambda\left[\Omega|\omega_j\rangle\right] = \omega_j\Lambda|\omega_j\rangle$$

Using the fact that Ω and Λ commute, we therefore have

$$\Omega\left[\Lambda|\omega_j\rangle\right] = \omega_j\Lambda|\omega_j\rangle$$

So, if $|\omega_j\rangle$ is an eigenvector of Ω, so is $\Lambda|\omega_j\rangle$. Assuming no degeneracies, then $\Lambda|\omega_j\rangle$ must be just a multiple of $|\omega_j\rangle$, $\Lambda|\omega_j\rangle = \lambda_j|\omega_j\rangle$, because the eigenvector $|\omega_j\rangle$ is specified completely up to a multiplicative constant when the eigenvalues of Ω are nondegenerate. But the statement $\Lambda|\omega_j\rangle = \lambda_j|\omega_j\rangle$ says that $|\omega_j\rangle$ is an eigenvector of Λ, too, with eigenvalue λ_j. Hence, the eigenvectors of Ω, which are orthonormal and provide a matrix representation in which Ω is diagonal, are also a set of eigenvectors for Λ and thus provide a basis in which its matrix representation is diagonal, too.

Note that the $\{\lambda_j\}$ may not necessarily form a nondegenerate set. The lack of degeneracy of Ω prevents this from being a problem.

Degenerate case:

Of course, as usual, one must think more in the degenerate case because, if ω_j is a degenerate eigenvalue, then the fact that $|\omega_j\rangle$ and $\Lambda|\omega_j\rangle$ are both eigenvectors of Ω with eigenvalue $|\omega_j\rangle$ does not imply $\Lambda|\omega_j\rangle = \lambda_j|\omega_j\rangle$; Λ could map $|\omega_j\rangle$ somewhere else in the degenerate subspace of eigenvectors of Ω with eigenvalue ω_j. It is straightforward to deal with this. Let us consider three cases:

- **Λ has degenerate eigenvalues but Ω does not.**
  As we stated above, this has no effect on the proof because it is nondegeneracy of Ω's eigenvectors that we relied on.
- **Ω has degenerate eigenvalues but Λ does not.**
  Simply exchange their roles — use Λ's nondegenerate eigenvectors as the diagonalizing basis.

The Eigenvector-Eigenvalue Problem: Formalism (cont.)

- **Both Ω and Λ have degenerate eigenvalues, with no correlation between which ones are degenerate.**
  Consider a degenerate subspace of Ω. Remember that one has complete freedom to pick a basis for the subspace — any basis will be orthogonal to all the other eigenvectors, and the subspace basis can always be made orthonormal using Gram-Schmidt. If Λ is not degenerate in this subspace, then simply use Λ's eigenvectors in the subspace. If Λ is only partially degenerate in the subspace, then break the subspace into subspaces using Λ's subspaces. Then the choice of basis for the residually degenerate subspaces is arbitrary and can always be made orthonormal.

  The same holds if Ω and Λ are equally degenerate in any given degenerate subspace — just create an orthonormal basis via Gram-Schmidt and it will be a perfectly good one.

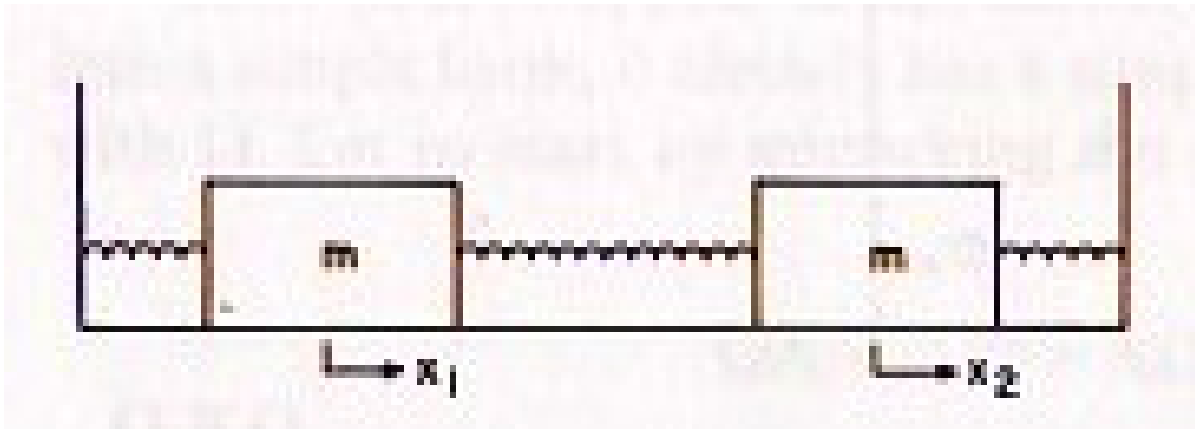
  The same holds in reverse, of course — if Λ has a degenerate subspace but Ω is nondegenerate or partially degenerate there, use Ω to further divide the subspace or to provide a basis for it.

Shankar has some discussion about the fact that matrices that are block diagonal in a subspace. It's not really necessary, as the discussion of block diagonal matrices assumes that one has already created a basis for the degenerate subspaces of Ω without consulting Λ. Even if one has, one can always pick a new, less degenerate one with Λ's help.

Example 3.32: Shankar's Normal Mode Example

Shankar's Example 1.8.6 does a normal mode problem in mechanics to demonstrate the entire above procedure. We do the same problem but a bit more explicitly so the various representations and operators involved are made more clear.

The problem consists of two masses coupled to each other by a spring and to fixed walls on either end. The position of each mass is measured relative to its rest position.



The Eigenvector-Eigenvalue Problem: A Worked Example (cont.)

Newton's Second Law for the system gives two coupled ordinary differential equations, which we may write in matrix form as

$$\begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{bmatrix} + \frac{k}{m} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

One assumes a harmonic solution with time dependence $e^{i\omega t}$ (of which we will take the real part in the end), so that

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} x_1(t=0) \\ x_2(t=0) \end{bmatrix} e^{i\omega t}$$

This form enables us to evaluate the time derivatives, leaving

$$\left(\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} - \lambda \right) \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} e^{i\omega t} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{3.85}$$

where $\lambda = \omega^2/(k/m)$.

The Eigenvector-Eigenvalue Problem: A Worked Example (cont.)

Let's translate the above equation into the language of inner product spaces and operators. The basis we begin with for this space, the one whose matrix representation we have used implicitly, is

$$|1\rangle \underset{|j\rangle}{\longleftrightarrow} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad |2\rangle \underset{|j\rangle}{\longleftrightarrow} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

(We subscript the $\leftrightarrow$ with the basis to show which matrix representation is being used.) We will denote this basis as the *coordinate basis* for this example because it is the basis in which the coefficients of the expansion of the state are just the coordinates of the two masses. An arbitrary vector (state for the system) is

$$|x\rangle = x_1|1\rangle + x_2|2\rangle \underset{|j\rangle}{\longleftrightarrow} x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

We define an operator Λ by its matrix representation in this basis, the 2×2 matrix in Equation 3.85:

$$\Lambda \underset{|j\rangle}{\longleftrightarrow} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

The Eigenvector-Eigenvalue Problem: A Worked Example (cont.)

The above equation is then of the form

$$(\Lambda - I\,\lambda)\,|x\,\rangle = |0\,\rangle \tag{3.86}$$

where divide out the $e^{i\omega t}$ factor because it never vanishes. Clearly, this is a characteristic equation of the form of Equation 3.61.

We proceed with the solution as outlined earlier. We find eigenvalues and eigenvectors:

$$\lambda_1 = 1 \qquad \omega_1 = \sqrt{\frac{k}{m}} \qquad |\lambda_1\,\rangle \underset{|j\,\rangle}{\longleftrightarrow} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\lambda_2 = 3 \qquad \omega_2 = \sqrt{3\,\frac{k}{m}} \qquad |\lambda_2\,\rangle \underset{|j\,\rangle}{\longleftrightarrow} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Since these eigenvectors and eigenvalues solve Equation 3.86, they solve Equation 3.85 and thus provide the two possible solutions to the problem,

$$|\lambda_1(t)\,\rangle = e^{i\,\omega_1 t}|\lambda_1\,\rangle \qquad |\lambda_2(t)\,\rangle = e^{i\,\omega_2 t}|\lambda_2\,\rangle$$

The Eigenvector-Eigenvalue Problem: A Worked Example (cont.)

A generic solution is thus of the form (representation-free and coordinate basis representation versions):

$$|x(t)\rangle = \mathcal{R}\left[c_1 e^{i\omega_1 t}\right]|\lambda_1\rangle + \mathcal{R}\left[c_2 e^{i\omega_2 t}\right]|\lambda_2\rangle \tag{3.87}$$

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \mathcal{R}\left[\frac{c_1}{\sqrt{2}}\, e^{i\omega_1 t}\right]\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \mathcal{R}\left[\frac{c_2}{\sqrt{2}}\, e^{i\omega_2 t}\right]\begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{3.88}$$

where $\mathcal{R}$ indicates "real part" (and $\mathcal{I}$ will indicate "imaginary part"). If we use the initial position and velocity conditions on the two masses, we find

$$c_1 = \frac{x_1(0)+x_2(0)}{\sqrt{2}} - \frac{i}{\omega_1}\frac{\dot{x}_1(0)+\dot{x}_2(0)}{\sqrt{2}} \qquad c_2 = \frac{x_1(0)-x_2(0)}{\sqrt{2}} - \frac{i}{\omega_2}\frac{\dot{x}_1(0)-\dot{x}_2(0)}{\sqrt{2}}$$

and thus, the full solution is

$$\begin{aligned}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \frac{1}{2}&\left\{[x_1(0)+x_2(0)]\cos\omega_1 t + \frac{1}{\omega_1}[\dot{x}_1(0)+\dot{x}_2(0)]\sin\omega_1 t\right\}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &+\left\{[x_1(0)-x_2(0)]\cos\omega_2 t + \frac{1}{\omega_2}[\dot{x}_1(0)-\dot{x}_2(0)]\sin\omega_2 t\right\}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\end{aligned} \tag{3.89}$$

Correct, but ugly and unilluminating.

The Eigenvector-Eigenvalue Problem: A Worked Example (cont.)

Let us now make more use of the machinery we have developed in order to obtain the same solution in a more elegant form that will generalize better to a larger vector space. In representation-free form, our solution was Equation 3.87, here written a bit more generically and also writing out the time derivative so we can make use of the initial conditions:

$$|x(t)\rangle = \sum_{j=1}^{2} \mathcal{R}\left[c_j e^{i\omega_j t}\right] |\lambda_j\rangle \qquad |\dot{x}(t)\rangle = \sum_{j=1}^{2} \mathcal{R}\left[i\,\omega_j c_j e^{i\omega_j t}\right] |\lambda_j\rangle$$

Let's relate the $\{c_j\}$ to the initial conditions in a more generic fashion than we did above. Take the inner product of both with $\langle\lambda_j|$ at $t = 0$, making use of the fact that $\langle\lambda_j\,|x(0)\rangle$ and $\langle\lambda_j\,|\dot{x}(0)\rangle$ are real by construction:

$$\langle\lambda_j\,|x(0)\rangle = \mathcal{R}[c_j] \qquad \langle\lambda_j\,|\dot{x}(0)\rangle = \mathcal{R}[i\,\omega_j c_j] \Leftrightarrow -\frac{1}{\omega_j}\langle\lambda_j\,|\dot{x}(0)\rangle = \mathcal{I}[c_j]$$

So, we may write our solution as

$$|x(t)\rangle = \sum_{j=1}^{2} \mathcal{R}\left\{\left[\langle\lambda_j\,|x(0)\rangle - \frac{i}{\omega_j}\langle\lambda_j\,|\dot{x}(0)\rangle\right] e^{i\omega_j t}\right\} |\lambda_j\rangle$$

The Eigenvector-Eigenvalue Problem: A Worked Example (cont.)

Let's reduce this by making use of the fact that both inner products $\langle \lambda_j \, | x(0) \, \rangle$ and $\langle \lambda_j \, | \dot{x}(0) \, \rangle$ are real, followed by rewriting in a suggestive form:

$$
\begin{aligned}
|x(t)\,\rangle &= \sum_{j=1}^{2} \left[\langle \lambda_j \, | x(0) \, \rangle \cos \omega_j t + \langle \lambda_j \, | \dot{x}(0) \, \rangle \, \omega_j^{-1} \sin \omega_j t \right] |\lambda_j\,\rangle \\
&= \left[\sum_{j=1}^{2} |\lambda_j\,\rangle\langle\lambda_j\,| \cos \omega_j t \right] |x(0)\,\rangle + \left[\sum_{j=1}^{2} |\lambda_j\,\rangle\langle\lambda_j\,| \, \omega_j^{-1} \sin \omega_j t \right] |\dot{x}(0)\,\rangle
\end{aligned}
$$

Define some operators:

$$
U(t) = \sum_{j=1}^{2} |\lambda_j\,\rangle\langle\lambda_j\,| e^{i\omega_j t} \qquad \Omega = \sum_{j=1}^{2} |\lambda_j\,\rangle\langle\lambda_j\,| \, \omega_j
$$

$$
\widetilde{U}(t) = \Omega^{-1} U(t) = \sum_{j,k=1}^{2} |\lambda_j\,\rangle\langle\lambda_j\,| \, \omega_j^{-1} e^{i\omega_j t}
$$

$$
U_R(t) = \mathcal{R}[U(t)] = \sum_{j=1}^{2} |\lambda_j\,\rangle\langle\lambda_j\,| \cos \omega_j t \qquad \widetilde{U}_I(t) = \mathcal{I}[\widetilde{U}(t)] = \sum_{j=1}^{2} |\lambda_j\,\rangle\langle\lambda_j\,| \, \omega_j^{-1} \sin \omega_j t
$$

The Eigenvector-Eigenvalue Problem: A Worked Example (cont.)

$U(t)$ is manifestly unitary. Ω is manifestly Hermitian. $U_R(t)$ is just the "real part" of $U(t)$ (where, by real part, we mean we take the real part of its diagonal matrix representation). $\widetilde{U}(t)$ is almost unitary; it would be if not for the Ω^{-1} factor. $\widetilde{U}_I(t)$ is its imaginary part. Using these new operators to rewrite, we have

$$|x(t)\rangle = U_R(t)|x(0)\rangle + \widetilde{U}_I(t)|\dot{x}(0)\rangle$$

So, we have a very simple expression for the time evolution of the state from its initial conditions. This result is complete — this fully specifies the time evolution.

However, we have not quite gotten to an expression that is as explicit as Equation 3.89 because we have $|x(t)\rangle$, $|x(0)\rangle$, and $|\dot{x}(0)\rangle$ rather than the matrix representations of these kets in the coordinate basis. We insert the identity operator and take the inner product with $\langle j|$:

$$\langle j|x(t)\rangle = \sum_{k=1}^{2}\langle j|U_R(t)|k\rangle\langle k|x(0)\rangle + \sum_{k=1}^{2}\langle j|\widetilde{U}_I(t)|k\rangle\langle k|\dot{x}(0)\rangle$$

where we inserted the $\{|k\rangle\}$ version to facilitate projecting the initial state onto the coordinate basis and took the inner product with $\langle j|$ to project the state at t onto the coordinate basis.

Written out in matrix representation, we thus have

$$\left[\ |x(t)\rangle\ \right]_{|j\rangle} = \left[\ U_R(t)\ \right]_{|j\rangle}\left[\ |x(0)\rangle\ \right]_{|j\rangle} + \left[\ \widetilde{U}_I(t)\ \right]_{|j\rangle}\left[\ |\dot{x}(0)\rangle\ \right]_{|j\rangle} \tag{3.90}$$

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \left[\ U_R(t)\ \right]_{|j\rangle}\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} + \left[\ \widetilde{U}_I(t)\ \right]_{|j\rangle}\begin{bmatrix} \dot{x}_1(0) \\ \dot{x}_2(0) \end{bmatrix} \tag{3.91}$$

The only thing missing here is that $U_R(t)$ and $\widetilde{U}_I(t)$ are diagonal in the $\{|\lambda_j\rangle\}$ matrix representation, but not in the $\{|j\rangle\}$ representation. To get between the two, we will need the unitary operator that transforms the $\{|j\rangle\}$ basis into the $\{|\lambda_j\rangle\}$ basis, which was defined in Equations 3.75 and 3.76:

$$R_{|j\rangle,|\lambda_j\rangle} \underset{|j\rangle}{\longleftrightarrow} \begin{bmatrix} \langle 1|\lambda_1\rangle & \langle 1|\lambda_2\rangle \\ \langle 2|\lambda_1\rangle & \langle 2|\lambda_2\rangle \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Note that $R_{|j\rangle,|\lambda_j\rangle} = R^{\dagger}_{|\lambda_j\rangle,|j\rangle}$.

With the unitary transformation operator in hand, let's use the arithmetic relation that lets us obtain the matrix representation in one basis from another, Equation 3.81, to obtain a more explicit expression:

$$
\begin{aligned}
\left[\ |x(t)\rangle\ \right]_{|j\rangle} = &\left[\ R^{\dagger}_{|\lambda_j\rangle,|j\rangle}\ \right]_{|\lambda_j\rangle}\left[\ U_R(t)\ \right]_{|\lambda_j\rangle}\left[\ R_{|\lambda_j\rangle,|j\rangle}\ \right]_{|\lambda_j\rangle}\left[\ |x(0)\rangle\ \right]_{|j\rangle} \\
&+\left[\ R^{\dagger}_{|\lambda_j\rangle,|j\rangle}\ \right]_{|\lambda_j\rangle}\left[\ \widetilde{U}_I(t)\ \right]_{|\lambda_j\rangle}\left[\ R_{|\lambda_j\rangle,|j\rangle}\ \right]_{|\lambda_j\rangle}\left[\ |\dot{x}(0)\rangle\ \right]_{|j\rangle} \\
= &\left[\ R_{|j\rangle,|\lambda_j\rangle}\ \right]_{|j\rangle}\left[\ U_R(t)\ \right]_{|\lambda_j\rangle}\left[\ R^{\dagger}_{|j\rangle,|\lambda_j\rangle}\ \right]_{|j\rangle}\left[\ |x(0)\rangle\ \right]_{|j\rangle} \\
&+\left[\ R_{|j\rangle,|\lambda_j\rangle}\ \right]_{|j\rangle}\left[\ \widetilde{U}_I(t)\ \right]_{|\lambda_j\rangle}\left[\ R^{\dagger}_{|j\rangle,|\lambda_j\rangle}\ \right]_{|j\rangle}\left[\ |\dot{x}(0)\rangle\ \right]_{|j\rangle}
\end{aligned}
\tag{3.92}
$$

where we used $R_{|j\rangle,|\lambda_j\rangle} = R^{\dagger}_{|\lambda_j\rangle,|j\rangle}$ and the fact that the matrix representations of $R_{|j\rangle,|\lambda_j\rangle}$ in the $\{|j\rangle\}$ and the $\{|\lambda_j\rangle\}$ bases are the same, Equation 3.84.

The Eigenvector-Eigenvalue Problem: A Worked Example (cont.)

Writing out more explicitly,

$$\begin{aligned}
&\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \qquad (3.93)\\
&= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \cos\omega_1 t & 0 \\ 0 & \cos\omega_2 t \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} \\
&\quad + \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \omega_1^{-1} & 0 \\ 0 & \omega_2^{-1} \end{bmatrix} \begin{bmatrix} \sin\omega_1 t & 0 \\ 0 & \sin\omega_2 t \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \dot{x}_1(0) \\ \dot{x}_2(0) \end{bmatrix}
\end{aligned}$$

which, you can check, is identical to Equation 3.89. It may look ugly in comparison, but it is much clearer about what is happening in obtaining the state at time t from the initial state. We start with the initial condition column vectors on the right side. Then we apply a unitary transformation operator to rewrite those initial conditions in terms of the eigenbasis of Λ. We then apply a diagonal time-evolution operator — this explicitly separates the time evolution of the two normal modes (eigenvectors). Then, we apply a unitary transformation to rewrite that final state in terms of the coordinate basis to get the desired final position of the masses. You can easily imagine how this can be generalized to an arbitrary number of masses n: all the matrices will grow to n dimensions, the unitary transformation matrix to go between the two bases will get more complicated, but the time evolution pieces will remain the same in form.

The Eigenvector-Eigenvalue Problem: A Worked Example (cont.)

The situation actually simplifies in QM. The Schrödinger Equation is a first-order differential equation in time, so we will only require something like the first term. The remainder will remain almost exactly the same. The product of the three square matrices in that term will be called the propagator because it propagates the initial conditions to a final state.

Generic Unitary Transformations

The unitary transformation operator that arose in the eigenvector-eigenvalue problem transforms from the original orthonormal basis $\{|j\rangle\}$ to the eigenbasis $\{|\omega_j\rangle\}$. It was initially defined by its matrix representation in the $\{|j\rangle\}$ basis, Equation 3.75

$$U_\Omega \leftrightarrow \begin{bmatrix} \langle 1|\omega_1\rangle & \cdots & \langle 1|\omega_n\rangle \\ \vdots & \ddots & \vdots \\ \langle n|\omega_1\rangle & \cdots & \langle n|\omega_n\rangle \end{bmatrix}$$

which we showed was equivalent to the representation-free statement given by Equation 3.76:

$$U_\Omega|j\rangle = |\omega_j\rangle \qquad U_\Omega^\dagger|\omega_j\rangle = |j\rangle$$

We could have instead simply used the first of the above pair above equations combined with the requirements of linearity and unitarity.

Unitary Transformations Revisited (cont.)

There was nothing about the above operator that was specific to the fact that $\{|\omega_j\rangle\}$ is an eigenbasis except for its effect on Ω. The definition of U_Ω simply required two orthonormal bases. Given any two orthonormal bases $\{|j\rangle\}$ and $\{|j'\rangle\}$, we can define a unitary operator in representation-free fashion by the generalization of Equation 3.76,

$$U_{|j\rangle,|j'\rangle}|j\rangle = |j'\rangle \qquad U^\dagger_{|j\rangle,|j'\rangle}|j'\rangle = |j\rangle \tag{3.94}$$

The subscripts denote the bases that U transforms from and to, in that order. The matrix elements of U in the two bases are

$$\begin{aligned}\langle j|U_{|j\rangle,|j'\rangle}|k\rangle &= \langle j|\left(U_{|j\rangle,|j'\rangle}|k\rangle\right) = \langle j|k'\rangle \\ \langle j'|U_{|j\rangle,|j'\rangle}|k'\rangle &= \left(U^\dagger_{|j\rangle,|j'\rangle}|j'\rangle\right)^\dagger|k'\rangle = \langle j|k'\rangle\end{aligned} \tag{3.95}$$

That is, the matrix representation of $U_{|j\rangle,|j'\rangle}$ is the same in the two bases and is the generalization of Equation 3.75,

$$U_{|j\rangle,|j'\rangle} \xleftrightarrow[|j'\rangle]{|j\rangle \text{ or}} \begin{bmatrix} \langle 1|1'\rangle & \cdots & \langle 1|n'\rangle \\ \vdots & \ddots & \vdots \\ \langle n|1'\rangle & \cdots & \langle n|n'\rangle \end{bmatrix} \tag{3.96}$$

Unitary Transformations Revisited (cont.)

As before in Equation 3.80, $U_{|j\,\rangle,|j'\,\rangle}$ transforms operators into new operators:

$$\Lambda' = U^{\dagger}_{|j\,\rangle,|j'\,\rangle}\ \Lambda\ U_{|j\,\rangle,|j'\,\rangle} \tag{3.97}$$

It also performs the equivalent of Equation 3.77, the arithmetic operation that converts the matrix representation of an operator in the $\{|j\,\rangle\}$ basis to its matrix representation in the $\{|j'\,\rangle\}$ basis

$$\left[\ \Lambda\ \right]_{|j'\,\rangle} = \left[\ U^{\dagger}_{|j\,\rangle,|j'\,\rangle}\ \right]_{|j\,\rangle} \left[\ \Lambda\ \right]_{|j\,\rangle} \left[\ U_{|j\,\rangle,|j'\,\rangle}\ \right]_{|j\,\rangle} \tag{3.98}$$

(the proof is identical to the one given leading up to Equation 3.77). And, as before in Equation 3.83, we have the confusing relation

$$\left[\ \Lambda'\ \right]_{|j\,\rangle} = \left[\ \Lambda\ \right]_{|j'\,\rangle} \tag{3.99}$$

that relates the matrix representation of the transformed operator in the untransformed basis to the matrix representation of the untransformed operator in the transformed basis.

Finally, it should be evident from Equation 3.94 that $U_{|j'\,\rangle,|j\,\rangle} = U^{\dagger}_{|j\,\rangle,|j'\,\rangle}$.

Functions of Operators

Does it make sense to consider functions of operators; *e.g.*, $f(\Omega) = e^{\Omega}$ where Ω is an operator?

Yes, it does, as long as we consider functions that can be written in terms of power series expansions. In such cases, if the power series expansion is

$$f(x) = \sum_{k=0}^{\infty} c_k x^k$$

then we simply make the obvious definition

$$f(\Omega) = \sum_{k=0}^{\infty} c_k \Omega^k \tag{3.100}$$

But, under what conditions does the expansion converge in the same way that the power series expansion converges?

To answer that question, we need to consider only operators that are Hermitian so we are certain they can be diagonalized. (Of course, the set of operators that can be diagonalized is larger, but we will only find it necessary in this course to consider Hermitian operators.) If we consider the operator in its eigenbasis, then it is diagonal. In that basis, Ω^n is given by taking the nth power of the diagonal, element-by-element. For example,

$$\left[\Omega^2\right]_{ij} = \sum_k \Omega_{ik}\Omega_{kj} = \sum_k \omega_i \delta_{ik} \omega_j \delta_{kj} = \omega_i \omega_j \delta_{ij} = \omega_i^2 \delta_{ij}$$

One can show in a similar way via inductive proof that the above property holds for Ω^n.

So, then, the expansion of Ω converges if the expansion converges for each eigenvalue when considered as a function of a number, not an operator; if it did not converge for some or all eigenvalues, some elements of the diagonal would be undefined.

Functions of Operators (cont.)

A typical example is simple exponentiation. In Ω's eigenbasis, we simply have

$$e^{\Omega} = \sum_{k=0}^{\infty} \frac{1}{k!}\Omega^k = \sum_{k=0}^{\infty} \frac{1}{k!} \begin{bmatrix} \omega_1 & 0 & \cdots & 0 \\ 0 & \omega_2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & \omega_n \end{bmatrix}^k$$

$$= \begin{bmatrix} \sum_{k=0}^{\infty} \frac{1}{k!}\omega_1^k & 0 & \cdots & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{1}{k!}\omega_2^k & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & \sum_{k=0}^{\infty} \frac{1}{k!}\omega_n^k \end{bmatrix}$$

$$= \begin{bmatrix} e^{\omega_1} & 0 & \cdots & 0 \\ 0 & e^{\omega_2} & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & e^{\omega_n} \end{bmatrix}$$

Related examples are sines and cosines and their hyperbolic counterparts.

When more than one operator is involved

The above examples went fairly easily because only one operator was involved. As soon as one starts working with expressions involving multiple operators, things begin to break down. A very simple example is exponentiation. Let's consider two expressions that would be equal if we considered numbers rather than operators:

$$
\begin{aligned}
e^{\alpha\Omega+\beta\Lambda} &= \sum_{k=0}^{\infty} \frac{1}{k!} (\alpha\Omega + \beta\Lambda)^k \\
&= I + (\alpha\Omega + \beta\Lambda) + \frac{1}{2} (\alpha\Omega + \beta\Lambda)^2 + \sum_{k=3}^{\infty} \frac{1}{k!} (\alpha\Omega + \beta\Lambda)^k \\
&= I + (\alpha\Omega + \beta\Lambda) + \frac{1}{2} \left[\alpha^2\Omega^2 + \alpha\beta (\Omega\Lambda + \Lambda\Omega) + \beta^2\Lambda^2\right] + \sum_{k=3}^{\infty} \frac{1}{k!} (\alpha\Omega + \beta\Lambda)^k \\
e^{\alpha\Omega} e^{\beta\Lambda} &= \left[\sum_{k=0}^{\infty} \frac{1}{k!} (\alpha\Omega)^k\right] \left[\sum_{m=0}^{\infty} \frac{1}{m!} (\beta\Lambda)^m\right]
\end{aligned}
$$

Functions of Operators (cont.)

Because the two expressions are equal for numbers, we know that it's just a matter, in the first expression, of moving all the Ωs to the left and the Λs to the right. But, if Ω and Λ do not commute, then that can't be done and the expressions are simply unequal **in general**.

Is this consistent with our statement about moving to the operator's eigenbasis to compute functions of operators?

Absolutely. We showed that if two operators commute, then they can be simultaneously diagonalized. If they can be simultaneously diagonalized, then the above exponential expressions can be evaluated for the diagonal elements in the eigenbasis and the two expressions will be equal. Conversely, though we did not show it, it is certainly true that two operators cannot be simultaneously diagonalized if they do not commute[1]. Hence, we would find that if we moved to the eigenbasis of one, say Ω, to compute its exponential from its diagonal elements, we would still not be able to commute Ω and Λ because Λ would be nondiagonal in Ω's eigenbasis.

[1] Proof by contradiction: suppose two operators that do not commute could be simultaneously diagonalized. Then there is a basis in which they are both diagonal. Diagonal matrices always commute. Whether two operators commute is independent of basis, so the two operators must commute in general. Contradiction.

How about differentiation?

Consider differentiation of an operator Θ whose elements are functions of a numerical parameter, λ. (No, we don't consider differentiation of an operator with respect to another operator!) The natural approach is to just write the standard definition of differentiation, replacing the function with the operator:

$$\frac{d}{d\lambda}\Theta(\lambda) = \lim_{\Delta\lambda\to 0} \frac{\Theta(\lambda+\Delta\lambda) - \Theta(\lambda)}{\Delta\lambda} \tag{3.101}$$

Since this operation is linear in the operator, the result is found by simple element-by-element differentiation of the matrix representing the operator:

$$\left[\frac{d}{d\lambda}\Theta(\lambda)\right]_{ij} = \frac{d}{d\lambda}\left[\Theta(\lambda)\right]_{ij} = \frac{d}{d\lambda}\Theta_{ij}(\lambda) \tag{3.102}$$

where the last two expressions are two different notations for the same thing. It may not always be possible rewrite this simply in terms of the original operator, but the algorithm is straightforward.

In some special cases, this simplifies. For example, consider exponentiation of a constant Hermitian operator with λ as a multiplying parameter, $e^{\lambda\Omega}$. We can calculate this in two ways: eigenbasis and power series.

- In the eigenbasis:

$$\left[\frac{d}{d\lambda}e^{\lambda\Omega}\right]_{ij} = \frac{d}{d\lambda}\left[e^{\lambda\Omega}\right]_{ij} = \frac{d}{d\lambda}e^{\lambda\omega_i}\delta_{ij} = \omega_i e^{\lambda\omega_i}\delta_{ij} = \left[\Omega e^{\lambda\Omega}\right]_{ij}$$

  where we were able to make the last notationally simplifying step only because of the particular form of the derivative of an exponential. Because this form is valid element-by-element in the eigenbasis, it therefore holds that

$$\frac{d}{d\lambda}e^{\lambda\Omega} = \Omega e^{\lambda\Omega}$$

  Of course, we could have placed Ω on the right side too since Ω and $e^{\lambda\Omega}$ commute.

Calculus with Operators (cont.)

- By power series:

$$\frac{d}{d\lambda}e^{\lambda\Omega} = \sum_{k=0}^{\infty}\frac{d}{d\lambda}\left[\frac{1}{k!}\lambda^k\Omega^k\right] = \sum_{k=1}^{\infty}\frac{1}{(k-1)!}\lambda^{k-1}\Omega^{k-1}$$
$$= \Omega\sum_{m=0}^{\infty}\frac{1}{m!}\lambda^m\Omega^m = \Omega e^{\lambda\Omega}$$

In either case, the process was simple because the dependence on λ was simple; Ω did not also depend on λ. It will in general be more complicated.

And integration?

Integration is also a linear operation, so, it can always be written as element-by-element integration as we did with differentiation:

$$\left[\int_{\lambda_0}^{\lambda} d\lambda'\,\Theta(\lambda')\right]_{ij} = \int_{\lambda_0}^{\lambda} d\lambda'\,\left[\Theta(\lambda')\right]_{ij} = \int_{\lambda_0}^{\lambda} d\lambda'\,\Theta_{ij}(\lambda') \tag{3.103}$$

where again the last two expressions are notationally equivalent. And, of course, in simple cases, such as the above exponentiation case, the result comes out cleanly and simply. Let's do the power series version:

$$\begin{aligned}\int_{\lambda_0}^{\lambda} d\lambda'\,\Omega\,e^{\lambda'\Omega} &= \sum_{k=0}^{\infty}\int_{\lambda_0}^{\lambda} d\lambda'\,\Omega\left[\frac{1}{k!}\left(\lambda'\right)^k\Omega^k\right] = \sum_{k=0}^{\infty}\frac{1}{(k+1)!}\left(\lambda^{k+1}-\lambda_0^{k+1}\right)\Omega^{k+1}\\ &= \sum_{m=0}^{\infty}\frac{1}{m!}\left(\lambda^m-\lambda_0^m\right)\Omega^m = e^{\lambda\Omega}-e^{\lambda_0\Omega}\end{aligned}$$

Calculus with Operators (cont.)

Note that we could only get the nice clean result with a perfect differential – if we had not put the Ω in front, we would have been missing a factor of Ω in the infinite sum. If Ω were invertible, we could have inserted a factor of $\Omega^{-1}\Omega = I$ and obtained

$$\int_{\lambda_0}^{\lambda} d\lambda' \, e^{\lambda' \Omega} = \Omega^{-1} \left(e^{\lambda \Omega} - e^{\lambda_0 \Omega} \right)$$

That's a special case, though.

Shankar summarizes the above examples by saying that, if one only has a single operator involved, then in general the standard expressions for numbers go through. We have added the caveat that one has to be sure that no division is necessary. In general, one must work through the power series expansion to be certain of doing things correctly.

Lecture 10:
Infinite-Dimensional Generalization

Date Given: 2008/10/22
Date Revised: 2008/10/22

Examples of Infinite-Dimensional Vector Spaces

Before getting into the business of how we generalize our previous work to infinite dimensions, let's first think of some examples. Vector spaces utilizing functions are the easiest way to obtain infinite-dimensional vector spaces. Three examples, seemingly similar but quite distinct:

- **All polynomials on the real line**
  The way to see the dimensionality of this space is to explicitly construct a basis. Let's denote each power of the argument of the polynomial x as a basis vector:

$$|n\rangle \longleftrightarrow x^n \tag{3.104}$$

  Any polynomial is just a finite linear combination of these basis vectors with real or complex coefficients (depending on the field we choose). Thus, we are assured the closure requirement is satisfied. The other arithmetic axioms follow quickly from the arithmetic properties of real or complex numbers. We thus have a basis set that is infinite, and hence the space's dimension is infinite. One important fine point is that space is infinite in the way that the integers are infinite – there is a "countably infinite" number of basis vectors.

- **All infinitely differentiable functions on the real line**
  By infinitely differentiable, we mean that the function can be differentiated as many times as one wants at any point and never yield nonsense (*i.e.*, infinity). It is fine for the derivatives of some order and higher to all vanish; such functions would be polynomials. But the vector space is much larger than the polynomials. One might be tempted to think that is is not: because of the differentiability requirement, any function in the set can be written as a countably infinite sum ($\sum_{i=0}^{\infty}$) of polynomials, and hence one might think it is only "countably infinite squared", which is just countably infinite. But consider that the sinusoids belong to this vector space. The period of the sinusoid can take on any real number value. These are all linearly independent because no sinusoids can be written in terms of other sinusoids when one considers the entire real line. So the number of sinusoids is at least a infinite as the set of real numbers. In a math class, one would prove that the set of real numbers is much more infinite than the set of integers.

- **All infinitely differentiable functions on the interval** $[0, 1]$
  The distinction between this space and the space of such functions on the entire real line is that this set is only countably infinite. You know from your study of Fourier series in Ph2/12 that any reasonably smooth function on an interval can be represented by a sum of sines and cosines. What is special about restricting to the interval is that one need only consider sinusoids that fit an integer number of periods in the interval to represent all functions on this interval. The sinusoids can be labeled with an integer (the number of periods) and whether they are sine or cosine. This set of functions is countably infinite and spans the entire space. Hence, the space is only countably infinite-dimensional.

Note that we have not yet attempted to define an inner product to make these inner product spaces. That is where much of the subtlety enters.

Infinite-Dimensional Generalization: From Finite to Infinite Dimensions

Functions Spaces on a Finite Number of Points Revisited

Recall our many uses of the example of complex-valued functions on n discrete points in the interval $[0, L]$, Examples 3.4 and 3.22 among them. We considered n points spaced out by $\Delta = L/(n+1)$, $x_j = j\,\Delta$ with $j = 1, \ldots, n$. Our vector space is the set of functions on this discrete set of points, $\{f(\{x_j\}\}$. We have a matrix representation in which each vector is represented by a column matrix consisting of the value of the function at the n points,

$$|f\rangle \leftrightarrow \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{bmatrix} \tag{3.105}$$

Infinite-Dimensional Generalization: From Finite to Infinite Dimensions (cont.)

The corresponding basis is the set of n functions that take on value 1 at the jth point and zero elsewhere:

$$|1\rangle \leftrightarrow \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad |2\rangle \leftrightarrow \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \qquad \cdots \qquad |n\rangle \leftrightarrow \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \tag{3.106}$$

Any vector f is simply written as

$$|f\rangle = \sum_{j=1}^{n} f(x_j)|j\rangle \tag{3.107}$$

The inner product is just the obvious matrix multiplication, which gives

$$\langle j|k\rangle = \delta_{jk} \tag{3.108}$$

Infinite-Dimensional Generalization: From Finite to Infinite Dimensions (cont.)

The above form for the inner product immediately tells us we can recover $f(x_j)$ using the inner product:

$$\langle j \,|f\rangle = \sum_{k=1}^{n} f(x_k)\langle j \,|k\rangle = f(x_j) \tag{3.109}$$

This is a particularly important relation, as it shows us how to recover the function from the abstract vector. Our orthonormality relation also tells us that the inner product of two arbitrary vectors is

$$\langle f \,|g\rangle = \sum_{j,k=1}^{n} f^*(x_j)\, g(x_k)\langle j \,|k\rangle = \sum_{j=1}^{n} f^*(x_j)\, g(x_j) \tag{3.110}$$

The norm of a vector is thus

$$\langle f \,|f\rangle = \sum_{j=1}^{n} \left|f(x_j)\right|^2 \tag{3.111}$$

Infinite-Dimensional Generalization: From Finite to Infinite Dimensions (cont.)

The identity operator may be represented in the standard fashion,

$$\sum_{j=1}^{n} |j\rangle\langle j| = I \tag{3.112}$$

This equality of the sum over all basis elements to the identity operator is also known as a completeness relation. It will prove important below.

Infinite-Dimensional Generalization: From Finite to Infinite Dimensions (cont.)

Now, we want to take the limit of $n \to \infty$. We need to redefine the inner product, though, to prevent it from becoming infinite:

$$\langle f \,|g\, \rangle = \sum_{j=1}^{n} f^*(x_j)\, g(x_j)\, \Delta \tag{3.113}$$

Now, as we let $n \to \infty$, we recognize that the sum converts to an integral:

$$\lim_{n\to\infty} \sum_{j=1}^{n} f^*(x_i)\, g(x_i)\Delta = \int_0^L dx\, f^*(x)\, g(x)$$

Note that the index has not just gone from finite to countably infinite; it is now as infinite as the real numbers. It makes no sense to talk about j anymore, we must now label the points on which the function is defined by their position x. For functions on an arbitrary interval $[a, b]$, we may generalize this to

$$\langle f \,|g\, \rangle = \int_a^b dx\, f^*(x)\, g(x)$$

Infinite-Dimensional Generalization: From Finite to Infinite Dimensions (cont.)

With the transition from the index j to position x, we also must sort out what happens to our basis kets. Consider their matrix elements. Clearly, we still require $\langle x\,|x'\rangle = 0$ for $x \neq x'$. To figure out what we need for $x = x'$, let's require that our completeness relation, Equation 3.112, still hold. It now takes the form

$$\lim_{n\to\infty} \sum_{j=1}^{n} \Delta\,|j\rangle\langle j| = \int_a^b dx'\,|x'\rangle\langle x'| = I \tag{3.114}$$

(We again use our standard method for converting the finite sum to an integral by inserting Δ. The change of dummy variable from x to x' facilitates the next step.) Let's apply this to $\langle x|$ on the left an an arbitary ket $|f\rangle$ on the right:

$$\int_a^b dx'\langle x|x'\rangle\langle x'|f\rangle = \langle x|\,I\,|f\rangle = \langle x|f\rangle$$

Infinite-Dimensional Generalization: From Finite to Infinite Dimensions (cont.)

We would certainly like to still have $\langle x\,|f\,\rangle = f(x)$ as we had in the finite case because we want to preserve our original finite-case definition, in which the coefficients of the expansion of the vector in orthonormal basis are the function values. Also, it provides us a means to relate the abstract vector to the function, which we must be able to do to define the vector in the first place! This requirement turns the above equation into a condition on $\langle x\,|x'\,\rangle$:

$$\int_a^b dx'\langle x\,|x'\,\rangle\, f(x') = f(x)$$

We shall rewrite the above condition using the fact that we have already required that $\langle x\,|x'\,\rangle = 0$ for $x \neq x'$ (and assuming $a < x < b$):

$$\begin{aligned} f(x) &= \int_a^b dx'\langle x\,|x'\,\rangle f(x') = \int_{x-\epsilon}^{x+\epsilon} dx'\langle x\,|x'\,\rangle f(x') \quad \text{for any } \epsilon > 0 \\ &= f(x)\int_{x-\epsilon}^{x+\epsilon} dx'\langle x\,|x'\,\rangle \quad \text{for any } \epsilon > 0 \end{aligned}$$

where the last step is possible assuming $f(x)$ is continuous at x.

Infinite-Dimensional Generalization: From Finite to Infinite Dimensions (cont.)

Since $f(x)$ was arbitrary, we therefore have a generic requirement on $\langle x\,|x'\,\rangle$:

$$1 = \int_{x-\epsilon}^{x+\epsilon} dx'\langle x\,|x'\,\rangle \quad \text{for any } \epsilon > 0 \tag{3.115}$$

which needs to be coupled with our orthogonality requirement

$$\langle x\,|x'\,\rangle = 0 \quad \text{for} \quad x \neq x'$$

We shall designate $\langle x\,|x'\,\rangle$ by $\delta(x - x')$ (because its value only depends on the difference $x - x'$) and refer to it as the Dirac delta function. We shall discuss its properties in detail below, but the above integral definition is really all we need.

Infinite-Dimensional Generalization: From Finite to Infinite Dimensions (cont.)

Motivated by the above, let's rigorously define our continuous x limit, pointing out how our earlier concepts of matrix representation apply:

- Our space consists of all the functions on the interval $[a, b]$. $|f\rangle$ designates the vector space member.
- We take as an orthonormal basis the kets $\{|x\rangle\}$ that are defined by

$$\langle x|x'\rangle = \delta(x - x') \quad \Longleftrightarrow \quad \begin{array}{l} \langle x|x'\rangle = 0 \quad \text{for} \quad x \neq x' \\ \int_{x-\epsilon}^{x+\epsilon} dx' \langle x|x'\rangle = 1 \text{ for any } \epsilon > 0 \end{array} \tag{3.116}$$

  The above implicitly defines the inner product for the space, also.
- We define an arbitrary vector $|f\rangle$ by its expansion in the $\{|x\rangle\}$ basis, which is given by the original function to which $|f\rangle$ corresponds:

$$f(x) \xleftrightarrow[|x\rangle]{} |f\rangle = \int_a^b dx\, f(x)\,|x\rangle = \lim_{n\to\infty} \sum_{j=1}^{n} \Delta\, f(x_j)\,|j\rangle \tag{3.117}$$

  (We have snuck in here the definition of an integral with $|x\rangle$ in the integrand, providing the limit to clarify what this means.) That is, the function $f(x)$ provides the expansion coefficients of the vector $|f\rangle$ in the $\{|x\rangle\}$ basis. Equivalently, **the function $f(x)$ is the matrix representation of the vector $|f\rangle$ in the infinite-dimensional $\{|x\rangle\}$ orthonormal basis**.

Infinite-Dimensional Generalization: From Finite to Infinite Dimensions (cont.)

- Our inner product rule $\langle x\,|x'\,\rangle = \delta(x-x')$ along with the expansion of $|f\,\rangle$ in this basis $|f\,\rangle = \int_a^b dx\, f(x)\,|x\,\rangle$ imply

$$\langle x\,|f\,\rangle = \int_a^b dx'\langle x\,|x'\,\rangle f(x') = f(x) \tag{3.118}$$

- As a result, the inner product between two members of the space is

$$\begin{aligned}\langle f\,|g\,\rangle &= \int_a^b dx \int_a^b dx'\, f^*(x)\langle x\,|x'\,\rangle\, g(x') \\ &= \int_a^b dx \int_a^b dx'\delta(x-x')\, f^*(x)\, g(x') = \int_a^b dx\, f^*(x)\, g(x)\end{aligned} \tag{3.119}$$

- The expansion assumption $|f\,\rangle = \int_a^b dx\, f(x)\,|x\,\rangle$ and the projection result $\langle x\,|f\,\rangle = f(x)$ imply

$$\langle x\,|f\,\rangle = \langle x\,|\int_a^b dx' f(x')\,|x'\,\rangle = \langle x\,|\int_a^b dx'|x'\,\rangle\langle x'\,|f\,\rangle = \langle x\,|\left[\int_a^b dx'|x'\,\rangle\langle x'\,|\right]|f\,\rangle$$

for arbitrary $|f\,\rangle$ and $\langle x\,|$, and thus we have the closure or completeness relation

$$\int_a^b dx\,|x\,\rangle\langle x\,| = I \tag{3.120}$$

Infinite-Dimensional Generalization: From Finite to Infinite Dimensions (cont.)

- Using the orthonormal basis definition, $\langle x\,|x'\,\rangle = \delta(x-x')$, the matrix representation assumption $|f\,\rangle \underset{|x\,\rangle}{\longleftrightarrow} f(x)$, and the projection result $f(x) = \langle x\,|f\,\rangle$, we may make a correspondence that defines $|x'\,\rangle$ explicitly:

$$|x'\,\rangle \underset{|x\,\rangle}{\longleftrightarrow} \delta(x-x') \tag{3.121}$$

Hopefully, given all of our discussion of matrix representations, especially of functions on discrete points, you are conceptually ready for the idea of the function $f(x)$ being the column-matrix representation in the $\{|x\,\rangle\}$ basis of the abstract vector $|f\,\rangle$, with $f(x)$ defining $|f\,\rangle$ by this representation. As we discussed in connection to finite-dimensional inner product spaces and their matrix representations, $|f\,\rangle$ is **not** the same thing as $f(x)$. $|f\,\rangle$ is an abstract object that belongs to the vector space. $f(x)$ is the component of $|f\,\rangle$ along the basis direction $|x\,\rangle$ and thus also gives the elements of the column matrix representation of $|f\,\rangle$ in the $\{|x\,\rangle\}$ basis. As before, confusion arises because we have to define $|f\,\rangle$ in some basis; we defined it in the $|x\,\rangle$ basis by saying $\langle x\,|f\,\rangle = f(x)$. But, as we will see later, there are other bases that we can decompose $|f\,\rangle$ in; of particular interest will be the $|p\,\rangle$ momentum basis, in which case $\langle p\,|f\,\rangle$ will be given by the Fourier transform of $f(x)$ and will tell us the projection of $|f\,\rangle$ onto a state of well-defined momentum, rather than position.

Infinite-Dimensional Generalization: Properties of the Dirac Delta Function

Getting to know the Dirac Delta Function

What is this thing $\delta(x - x')$? Intuitively, it is "something" that vanishes everywhere except when its argument vanishes, at which point its value must become infinite to make the integral in Equation 3.115 nonzero.

You should be offended by such an object; to help you accept this function, think of it as one of the following limits of a reasonable function:

$$\delta(x) = \lim_{\Delta \to 0} \begin{cases} \frac{1}{\Delta} & |x| < \frac{\Delta}{2} \\ 0 & |x| \geq \frac{\Delta}{2} \end{cases} \qquad \delta(x) = \lim_{\Delta \to 0} \frac{1}{\sqrt{\pi\,\Delta^2}} \exp\left(-\frac{x^2}{\Delta^2}\right) \tag{3.122}$$

However, any derivation involving delta functions must first and foremost rely only on properties derivable from the defining integral condition

$$\int_{x-\epsilon}^{x+\epsilon} dx'\,\delta(x - x')f(x') = f(x) \tag{3.123}$$

If the above limits worked in all cases, we would just use them to define the delta function! In particular, when manipulating delta functions, one must ensure all steps are justified by the integral definition. The typical mistake is to assume without proof that the delta function obeys the standard rules for functions, derivatives, or integrals.

Infinite-Dimensional Generalization: Properties of the Dirac Delta Function (cont.)

Let's perform some basic manipulations both to give you some experience with delta functions as well as to derive some useful results.

Let's first look at derivatives. Here, we will in general look at the delta function as a function of two variables, x and x':

$$\widetilde{\delta}(x, x') = \delta(x - x')$$

The fact that we write the argument as $x - x'$ implies that the function really depends only on the difference between the two arguments. But, since we will usually integrate over one of the two arguments, it is useful to think of it as a function of two arguments, also.

Infinite-Dimensional Generalization: Properties of the Dirac Delta Function (cont.)

The derivative with respect to the delta function's first argument is obtained by using the general definition of the derivative and careful manipulation of the integral condition:

$$\int_{x-\epsilon}^{x+\epsilon} dx' \left[\frac{d}{dx}\,\delta(x-x')\right] f(x') = \int_{x-\epsilon}^{x+\epsilon} dx' \left[\lim_{\alpha\to 0}\frac{\delta(x+\alpha-x')-\delta(x-x')}{\alpha}\right] f(x')$$

$$= \lim_{\alpha\to 0}\frac{1}{\alpha}\left[\int_{x-\epsilon}^{x+\epsilon} dx'\delta(x-(x'-\alpha))\,f(x') - \int_{x-\epsilon}^{x+\epsilon} dx'\delta(x-x')\,f(x')\right]$$

$$= \lim_{\alpha\to 0}\frac{1}{\alpha}\left[\int_{x-\epsilon}^{x+\epsilon} dx'\delta(x-x')\,f(x'+\alpha) - \int_{x-\epsilon}^{x+\epsilon} dx'\delta(x-x')\,f(x')\right]$$

$$= \int_{x-\epsilon}^{x+\epsilon} dx'\delta(x-x') \lim_{\alpha\to 0}\frac{f(x'+\alpha)-f(x')}{\alpha}$$

$$= \int_{x-\epsilon}^{x+\epsilon} dx'\delta(x-x')\,\frac{d}{dx'}\,f(x') = \frac{d}{dx}\,f(x)$$

where we have assumed, since we will take $\alpha \to 0$ in the end, that $|\alpha| \ll \epsilon$ so that there is no worry about $x+\alpha$ moving outside the limits of the integral; thus we may move the limit inside and outside the integral. In going from the third to the fourth expression, we did a change of variables in the first integral, replacing $x'-\alpha$ by x' and thus changing $f(x')$ to $f(x'+\alpha)$.

Infinite-Dimensional Generalization: Properties of the Dirac Delta Function (cont.)

An alternative method is the following:

$$\int_{x-\epsilon}^{x+\epsilon} dx' \left[\frac{d}{dx}\,\delta(x-x')\right] f(x') = \frac{d}{dx}\int_{x-\epsilon}^{x+\epsilon} dx'\,\delta(x-x')\,f(x') = \frac{d}{dx}\,f(x)$$

(It was ok to pull the derivative outside the integral because the endpoints of the integral don't really depend on x; we just require that x be in the integration interval.) We will see next that this second method is less generalizable.

In either case, we may write the result concisely in the following form:

$$\frac{d}{dx}\,\delta(x-x') = \delta(x-x')\frac{d}{dx'} \tag{3.124}$$

where d/dx' is to act on any functions to its right that depend on x' and it is understood that this form only makes sense when integrated with a function of x':

$$\int_{x-\epsilon}^{x+\epsilon} dx' \left[\frac{d}{dx}\,\delta(x-x')\right] f(x') = \int_{x-\epsilon}^{x+\epsilon} dx'\,\delta(x-x')\frac{d}{dx'}\,f(x') = \frac{d}{dx}\,f(x) \tag{3.125}$$

where we treat $\frac{d}{dx}\,f(x)$ like any other function when acted on by the delta function.

Lecture 11:
Infinite-Dimensional Generalization continued

Date Given: 2008/10/24
Date Revised: 2008/10/24

Infinite-Dimensional Generalization: Properties of the Dirac Delta Function

What about the derivative with respect to the second argument, $\frac{d}{dx'}\,\delta(x-x')$? Along the lines of the second proof above, one might think one can figure this out by integration by parts:

$$\int_{x-\epsilon}^{x+\epsilon} dx' \left[\frac{d}{dx'}\,\delta(x-x')\right] f(x')$$

$$= \left[\delta(x-x')\,f(x')\right]\Big|_{x-\epsilon}^{x+\epsilon} - \int_{x-\epsilon}^{x+\epsilon} dx'\delta(x-x')\,\frac{d}{dx'}\,f(x') = -\frac{d}{dx}\,f(x)$$

where the first term from the integration by parts vanished because the delta function is zero at the endpoints of the interval (or, in the limit that the interval goes to zero, the values at the two endpoints become equal and infinite and so the term vanishes) and we simply integrated the second term using the usual properties of the delta function. However, in the above, we have implicitly assumed

$$\frac{d}{dx'}\left[\delta(x-x')f(x')\right] = \left[\frac{d}{dx'}\,\delta(x-x')\right] f(x') + \delta(x-x')\frac{d}{dx'}\,f(x')$$

which is a statement we have never justified.

Infinite-Dimensional Generalization: Properties of the Dirac Delta Function (cont.)

Instead, let us calculate this derivative by a less suspect procedure that is more like our first proof for the derivative with respect to the first argument:

$$\int_{x-\epsilon}^{x+\epsilon} dx' \left[\frac{d}{dx}\,\delta(x-x')\right] f(x') = \int_{x-\epsilon}^{x+\epsilon} dx' \left[\lim_{\alpha\to 0} \frac{\delta(x-(x'+\alpha)) - \delta(x-x')}{\alpha}\right] f(x')$$

The above expression is identical to the third expression in the corresponding proof up to a minus sign, so we may use the remainder of that proof:

$$\int_{x-\epsilon}^{x+\epsilon} dx' \left[\frac{d}{dx'}\,\delta(x-x')\right] f(x') = -\frac{d}{dx}\,f(x)$$

Thus, we find

$$\frac{d}{dx'}\,\delta(x-x') = -\delta(x-x')\,\frac{d}{dx'} = -\frac{d}{dx}\,\delta(x-x') \tag{3.126}$$

The change of sign between Equations 3.124 and 3.126 makes sense because of the sign difference between x and x' in the argument of the δ function: it is a matter of whether one is taking a derivative in the usual "right-going" direction or in the opposite "left-going" direction.

Infinite-Dimensional Generalization: Properties of the Dirac Delta Function (cont.)

One can show by inductive proof that the action of an arbitrary-order derivative of the delta function is similar:

$$\frac{d^n}{dx^n}\delta(x-x') = \delta(x-x')\frac{d^n}{dx'^n} \tag{3.127}$$

Infinite-Dimensional Generalization: The X Operator and the $\{|x\rangle\}$ Basis

Is there an operator for which the $\{|x\rangle\}$ basis is an eigenbasis?

We have gone about defining the $\{|x\rangle\}$ basis in a rather backwards fashion: Rather than first defining a Hermitian operator on the inner product space, solving for its eigenvalues and eigenvectors, and then choosing the eigenbasis of the operator as a nice basis for our space, we just defined the basis in the finite-dimensional case and extended it to infinite dimensions. We have made the definition of $|x\rangle$ as explicit as we can in that we have made the correspondence (Equation 3.121)

$$|x'\rangle \xleftrightarrow[|x\rangle]{} \delta(x - x')$$

But, is there an operator whose eigenbasis is the $\{|x\rangle\}$ basis?

Infinite-Dimensional Generalization: The X Operator and the $\{|x\rangle\}$ Basis (cont.)

Yes: let's just define an operator X by defining its action on the $\{|x\rangle\}$ basis:

$$X|x\rangle = x|x\rangle \tag{3.128}$$

where x is the position along the interval to which $|x\rangle$ corresponds. In the finite N case, this operator would have been defined by the relation

$$X|j\rangle = j\Delta\,|j\rangle$$

We have defined the operator X so that its eigenbasis is $\{|x\rangle\}$. This may be confusing: x is a function of x, so does it make any sense for x to be an eigenvalue? While x is a function, it is also just a label for directions in the inner product space, for members of a basis for the space. So, we are saying that the eigenvalue for the eigenvector $|x\rangle$ is just related to its label. This is evident from the analogous definition for the finite N case.

Infinite-Dimensional Generalization: The X Operator and the $\{|x\rangle\}$ Basis (cont.)

Let's pursue the implications of our definition of X. Based on our definition of $\langle x' | x \rangle$, this operator's matrix elements are

$$X_{xx'} \equiv \langle x | X | x' \rangle = x' \langle x | x' \rangle = x' \delta(x - x') = x \delta(x - x') \tag{3.129}$$

(We introduce the $_{xx'}$ labeling for the matrix elements of an operator between that states $|x\rangle$ and $|x'\rangle$.) What is the action of X on some arbitrary ket $|f\rangle$? Define

$$|g\rangle = X |f\rangle$$

Let's expand $|g\rangle$ in our $|x\rangle$ basis:

$$\langle x | g \rangle = \langle x | X | f \rangle = \int_a^b dx' \langle x | X | x' \rangle \langle x' | f \rangle = \int_a^b dx' x' \delta(x - x') f(x') = x f(x) \tag{3.130}$$

So, $|g\rangle = |x f\rangle$ or $g(x) = x f(x)$.

The Continuous Derivative Operator

We know that taking derivatives converts functions to other functions, so we expect the action of taking a derivative to be an operator on our vector space of functions. We also expect we can construct a derivative operator by extending the discrete derivative operator from Example 3.22, which we do by taking the usual limit of $\Delta \to 0$. Explicitly, let's consider the projection of $D_R|f\rangle$ onto $\langle j|$:

$$\lim_{\Delta\to 0}\langle j|D_R|f\rangle = \lim_{\Delta\to 0}\frac{f(x_{j+1})-f(x_j)}{\Delta} = \lim_{\Delta\to 0}\frac{f(x+\Delta)-f(x)}{\Delta} = \frac{df}{dx}$$

We also know that $\lim_{\Delta\to 0}\langle j| = \langle x|$, so we have (dropping the $_R$ subscript because it is not necessary after the limit is taken):

$$\langle x|D|f\rangle = \left\langle x\middle|\frac{df}{dx}\right\rangle = \frac{df}{dx} \tag{3.131}$$

The above equation defines the action of the derivative operator because it tells us the projection of $D|f\rangle$ onto every basis element $|x\rangle$ for any vector $|f\rangle$.

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

Can we obtain the matrix elements of D in the $\{|x\rangle\}$ basis? Yes, we may infer them using the above expression. First, let's begin with our usual trick of inserting an identity operator in order to get an expression involving the matrix elements of D:

$$\langle x \,|\, D \,|f\rangle = \int_a^b dx' \langle x \,|\, D \,|x'\rangle \langle x' \,|f\rangle = \int_a^b dx' \langle x \,|\, D \,|x'\rangle f(x') \tag{3.132}$$

Next, let's use Equation 3.131 to replace $\langle x \,|D|f\rangle$ on the left:

$$\frac{d}{dx} f = \int_a^b dx' \langle x \,|\, D \,|x'\rangle f(x') \tag{3.133}$$

By comparison to Equation 3.125, we see that $\langle x \,|\, D \,|x'\rangle$ has the same behavior as $\left[\frac{d}{dx}\delta(x-x')\right] = \delta(x-x')\frac{d}{dx'}$; *i.e.*,

$$D_{xx'} \equiv \langle x \,|\, D \,|x'\rangle = \left[\frac{d}{dx}\delta(x-x')\right] = \delta(x-x')\frac{d}{dx'} \tag{3.134}$$

where, again, we index the matrix elements of the operator by x and x' as we did for the X operator in Equation 3.129.

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

It is interesting to see that we could have derived the same expression from the matrix element of D_R between $\langle j\,|$ and $|k\,\rangle$ from Example 3.22, which was

$$\langle j\,|D_R|k\,\rangle = -\frac{\langle j\,|k\,\rangle - \langle j\,|k-1\,\rangle}{\Delta}$$

Letting $x = j\Delta$ and $x' = k\Delta$, we can take the limit of $\Delta \to 0$:

$$\begin{aligned}\langle x\,|D|x'\,\rangle &= \lim_{\Delta\to 0}\langle j\,|D_R|k\,\rangle = \lim_{\Delta\to 0} -\frac{\langle j\,|k\,\rangle - \langle j\,|k-1\,\rangle}{\Delta} = \lim_{\Delta\to 0} -\frac{\langle x\,|x'\,\rangle - \langle x\,|x'-\Delta\,\rangle}{\Delta} \\ &= \lim_{\Delta\to 0}\frac{\delta(x-(x'-\Delta)) - \delta(x-x')}{\Delta} = \lim_{\Delta\to 0}\frac{\delta((x+\Delta)-x') - \delta(x-x')}{\Delta} \\ &= \frac{d}{dx}\delta(x-x')\end{aligned}$$

We prefer the first method for proving this, though, because it uses the defining integral properties of the delta function and requires a bit less guesswork in taking the limit. For example, relating $|x\,\rangle$ to $|j\,\rangle$ is a bit dangerous because there is a units problem: $\langle j\,|k\,\rangle = \delta_{jk}$ is a manifestly unitless quantity, while $\langle x\,|x'\,\rangle = \delta(x-x')$ has units of length^{-1} because it gives 1 when integrated over x'.

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

Making the Derivative Operator Hermitian

While we have been able to define the derivative operator D, we will see below that it is not Hermitian: this means that its eigenvalues need not be real numbers, which of course we would like so we can obtain an observable from it. We need to fix this.

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

Let's first calculate the matrix elements of $D^\dagger$ so we can see how D is not Hermitian. The rigorous way to do this, where we do not have to make somewhat questionable manipulations on delta functions, is to work with $\langle f\,|D^\dagger|x'\,\rangle$. (The rationale for using $|x'\,\rangle$ will become clear.) We know

$$\langle f\,|D^\dagger|x'\,\rangle = \left[\langle x'\,|D|f\,\rangle\right]^* = \left[\frac{d}{dx'}\,f\right]^*$$

Let's again use the expansion of $|f\,\rangle$, though:

$$\langle f\,|D^\dagger|x'\,\rangle = \int_{x'-\epsilon}^{x'+\epsilon} dx\,[f(x)]^*\langle x\,|D^\dagger|x'\,\rangle$$

Equating the two, we have

$$\int_{x'-\epsilon}^{x'+\epsilon} dx\,\langle x\,|D^\dagger|x'\,\rangle\,[f(x)]^* = \left[\frac{d}{dx'}\,f\right]^*$$

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

The above expression corresponds directly to Equation 3.124 with a simple exchange of x and x'; that's of course allowed since they are just dummy variables. (Do not let the presence of $[f(x)]^*$ and $[df/dx']^*$ confuse you; $[f(x)]^*$ is just an arbitrary function and hence the representation of an arbitrary element of the vector space.) So we have

$$\langle x|D^\dagger|x'\rangle = \frac{d}{dx'}\delta(x'-x)$$

Therefore

$$\begin{aligned}[D^\dagger]_{xx'} = \langle x|D^\dagger|x'\rangle = \frac{d}{dx'}\delta(x'-x) = \frac{d}{dx'}\delta(x-x') = -\frac{d}{dx}\delta(x-x') \\ = -\langle x|D|x'\rangle = -D_{xx'} \end{aligned} \tag{3.135}$$

where the first step is simply the definition of matrix element (recalling that x and x' are just indices now), the second step uses our result for the matrix element of $D^\dagger$, the third uses the evenness of the delta function, and the fourth uses Equations 3.124 and 3.126 together.

Recall that the $\{|x\rangle\}$ are a basis for the space and the above holds for any x and x'. Therefore, all the matrix elements in the $\{|x\rangle\}$ basis of $D^\dagger$ and $-D$ are equal; this implies $D^\dagger = -D$, and thus D is in fact anti-Hermitian instead of Hermitian!

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

The obvious solution is to consider a new operator K with

$$K = -iD \tag{3.136}$$

(The reason for the negative sign will become apparent later.) The Hermiticity requirement seems obviously met because the $-i$ provides the necessary sign flip. However, we must be careful about believing the above arithmetic – recall that these expressions only hold true when included in an integral. If we consider the expression $\langle g \,|\, K \,|f\,\rangle$, we see that this caveat becomes apparent. We first note that, if K is Hermitian, we have

$$\langle g \,|\, K \,|f\,\rangle = \langle g \,|\, [K|f\,\rangle] = \left\{ [K|f\,\rangle]^\dagger \,|g\,\rangle \right\}^* = \langle f \,|\, K^\dagger |g\,\rangle^* = \langle f \,|\, K \,|g\,\rangle^* \tag{3.137}$$

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

Let's calculate the expressions on the two ends explicitly by going through the matrix elements in the $\{|x\rangle\}$ basis:

$$\langle g \,|\, K \,|f\rangle = \int_a^b dx \int_a^b dx' \langle g \,|x\rangle\langle x \,|\, K \,|x'\rangle\langle x' \,|f\rangle \tag{3.138}$$

$$= \int_a^b dx\, g^*(x) \left[-i\frac{df}{dx}\right] = -i\int_a^b dx\, g^*(x) \left[\frac{df}{dx}\right]$$

$$\langle f \,|\, K \,|g\rangle^* = \left[\int_a^b dx \int_a^b dx' \langle f \,|x\rangle\langle x \,|\, K \,|x'\rangle\langle x' \,|g\rangle\right]^* \tag{3.139}$$

$$= \left[\int_a^b dx\, f^*(x) \left[-i\frac{dg}{dx}\right]\right]^* = i\int_a^b dx \left[\frac{dg^*}{dx}\right] f(x)$$

These two expressions are equal via integration by parts only if the surface term vanishes:

$$-i\, g^*(x)\, f(x)\Big|_a^b \tag{3.140}$$

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

Thus, **in order to make K a Hermitian operator, we must restrict our vector space to contain only functions that meet the condition that the above surface term vanishes for all members of the space.** This restricted space is called the physical Hilbert space, the qualifier *physical* included to distinguish it from the mathematical definition of a Hilbert space.

Shankar gives the example that this condition might be met by functions that vanish at the endpoints. Another example would be functions that take on equal values at the two endpoints. Shankar discusses a couple of other cases. It suffices here to say that conditions are frequently placed on the functions that can belong to the vector space of states in order to ensure that desired Hermitian operators are Hermitian.

Lecture 12:
Infinite-Dimensional Generalization continued

Date Given: 2008/10/27
Date Revised: 2008/10/27

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis

Eigenvalues and Eigenvectors of K

We have defined a Hermitian operator, so, assuming our finite-dimensional theorems continue to hold, we expect eigenvalues and eigenvectors. Let us find them.

Since our description so far of our vector space and operators has been in the $\{|x\rangle\}$ basis, we need to work through that basis to find the eigenvalues of K. This is the equivalent of defining a vector space by its representation in terms of one basis (say, $\{|j\rangle\}$) and then being given an operator Ω and wanting to find the eigenvalues and eigenvectors of Ω. One always needs to write down a representation to find eigenvalues and eigenvectors (there is no way to calculate the determinant, otherwise!). Since one only has the $\{jket\}$ basis and its matrix representation at this point, one must do this in that representation.

Let us denote the eigenvalues and eigenvectors as $\{k\}$ and $\{|k\rangle\}$. We of course require

$$K|k\rangle = k|k\rangle \tag{3.141}$$

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

Let's look at the matrix element of this ket with the $\{|x\rangle\}$ basis:

$$\begin{aligned}
\langle x|K|k\rangle &= k\langle x|k\rangle \\
\int_a^b dx'\langle x|K|x'\rangle\langle x'|k\rangle &= k\,\psi_k(x) \\
-i\int_a^b dx'\delta(x-x')\frac{d}{dx'}\psi_k(x) &= k\,\psi_k(x) \\
-i\frac{d}{dx}\psi_k(x) &= k\,\psi_k(x) \qquad (3.142)
\end{aligned}$$

where we have defined $\psi_k(x) = \langle x|k\rangle$ to be the $\{|x\rangle\}$ basis representation of $|k\rangle$ and used the known matrix elements of K in the $\{|x\rangle\}$ basis.

Equation 3.142 is a simple differential equation defining $\psi_k(x) = \langle x|k\rangle$; the solution is

$$\langle x|k\rangle = \psi_k(x) = A\,e^{ikx} \qquad (3.143)$$

where k is a constant and A is the unspecified normalization. The allowed values of k and the normalization depend now on the integration limits. We consider two cases (Shankar seems to only consider the second):

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

- **Finite interval** $[a, b]$
  Recall our condition that the coordinate representation $f(x)$ of any vector $|f\rangle$ in the space be equal at the endpoints in order for K to be Hermitian. Let's take $L = b - a$. The condition on the endpoints means that we may only consider solutions that satisfy

$$\psi_k(a) = \psi_k(a + L)$$
$$A\, e^{i\,k\,a} = A\, e^{i\,k\,a} e^{i\,k\,L}$$
$$k\,L = 2\,\pi\,j \quad \Longleftrightarrow \quad k = \frac{2\,\pi}{L}\,j \qquad \text{for } j \text{ any integer}$$

  Note that we explicitly find that k must be real in order for this condition to be met; otherwise, one gets a $e^{-\mathcal{I}(k)L}$ factor that explicitly violates the condition. k is discretized, though it has (countably) infinitely many allowed values. The natural normalization is $A = \sqrt{1/L}$ so that $\langle k\,|k\,\rangle = 1$, so we have

$$\langle x\,|k_j\,\rangle = \psi_{k_j}(x) = \sqrt{\frac{1}{L}}\, e^{ik_j x} \qquad k_j = \frac{2\,\pi}{L}\,j \qquad j = \text{any integer} \qquad (3.144)$$
$$\langle k_j\,|k_m\,\rangle = \delta_{jm} \equiv \delta_{k_j k_m}$$

  where the restriction to real, discretized k defines the *physical Hilbert space*.

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

- **Infinite interval** $(-\infty, \infty)$
  Do we have any condition on k in this case? We must still ensure the surface term is eliminated in order to make K Hermitian. It is problematic in that its value is ill-defined: e^{ikx} takes on no single value as $|x| \to \infty$. Shankar offers a rather dubious method of dealing with this. A slightly better, but still mathematically unrigorous, solution is to insert a converging factor $e^{-\beta|x|}$, do the calculation, and let $\beta \to 0$. In that case, the surface term is

$$\begin{aligned} e^{i(k-k')x}\Big|_{-\infty}^{\infty} &= \left[\lim_{\beta\to 0} e^{-\beta|x|} e^{-\mathcal{I}(k-k')x} e^{i\mathcal{R}(k-k')x}\right]\Big|_{-\infty}^{\infty} \\ &= \lim_{\beta\to 0}\left[e^{-\beta|x|} e^{-\mathcal{I}(k-k')x} e^{i\mathcal{R}(k-k')x}\right]\Big|_{-\infty}^{\infty} \\ &= \lim_{\beta\to 0} 0 = 0 \qquad \text{if } \mathcal{I}(k-k') = 0 \end{aligned}$$

  (Exchanging the limit in β and the implied limit of the endpoints going to $\pm\infty$ is not a mathematically rigorous procedure.) We had to require $\mathcal{I}(k-k') = 0$ so that the $e^{-\mathcal{I}(k-k')x}$ term would not cause the expression to diverge at one of $x = \pm\infty$ as $\beta \to 0$. This restriction defines the *physical Hilbert space* in this case.

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

Because normalization to a finite number is no longer sensible, we use the same normalization condition as was found for the $\{|x\rangle\}$ basis, normalization to a δ function, $\langle k\,|k'\rangle = \delta(k-k')$. However, this does not specify A:

$$\langle k\,|k'\rangle = \int_{-\infty}^{\infty} dx \langle k\,|x\rangle\langle x\,|k'\rangle = |A|^2 \int_{-\infty}^{\infty} dx\, e^{i(k'-k)x}$$

If $\langle k\,|k'\rangle = \delta(k-k')$, then the left side is either infinite or vanishing, which only tells us the obvious fact that $A = 0$ is not allowed. No other information about $|A|$ is provided.

To obtain $|A|$, we need to take the limit of the finite interval case. Certainly, we need for our $\{|k\rangle\}$ bases in the two cases to be complete, implying that the sum over all basis elements should be the identity operator. That is, we expected

$$I_L = \sum_{j=-\infty}^{\infty} |k_j\rangle\langle k_j| \qquad I_\infty = \int_{-\infty}^{\infty} dk |k\rangle\langle k|$$

where $_L$ and $_\infty$ indicate the two different spaces. Note that the $|k_j\rangle$ and $|k\rangle$ are different because one is for a finite interval and the other for an infinite interval.

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

Taking the inner product with $\langle x\,|$ and $|x'\rangle$, we have

$$\langle x\,|x'\rangle = \sum_{j=-\infty}^{\infty} \langle x\,|k_j\rangle\langle k_j\,|x'\rangle \qquad \langle x\,|x'\rangle = \int_{-\infty}^{\infty} dk \langle x\,|k\rangle\langle k\,|x'\rangle$$

$$\delta(x-x') = \frac{1}{L}\sum_{j=-\infty}^{\infty} e^{-ikx}e^{ikx'} \qquad \delta(x-x') = |A|^2 \int_{-\infty}^{\infty} dk\, e^{-ikx}e^{ikx'}$$

We see that we obtain delta functions in $x-x'$ for both expressions. Though the vectors $|k_j\rangle$ and $|k\rangle$ do not live in the same space, a function is just a set of numbers and we may compare functions at particular values of their arguments. The only hitch is that x and x' must be inside $[a,b]$ for the finite interval case but may take on any value for the infinite interval case. So, let's let $a \longrightarrow -\infty$, $b \longrightarrow \infty$ so that $L \longrightarrow \infty$. Equating the two sides in this limit, we have

$$\lim_{L\to\infty} \frac{1}{L}\sum_{j=-\infty}^{\infty} e^{ik_j(x'-x)} = |A|^2 \int_{-\infty}^{\infty} dk\, e^{ik(x'-x)}$$

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

For the interval case, the allowed k values are spaced by $\Delta k = 2\pi/L$. Let's rewrite using Δk:

$$\frac{1}{2\pi} \lim_{L\to\infty} \sum_{j=-\infty}^{\infty} \Delta k \, e^{ik_j(x'-x)} = |A|^2 \int_{-\infty}^{\infty} dk \, e^{ik(x'-x)}$$

The sum on the left side is the same as the integral on the right side when one takes the limit $L \to \infty$ because $\Delta k \to 0$ and k_j becomes a continuous variable in that limit. But we see the two sides are equal if and only if $|A|^2 = 1/2\pi$. The phase of A is arbitrary, so we choose it to be real. In summary, then, we have

$$\langle x\,|k\,\rangle = \psi_k(x) = \sqrt{\frac{1}{2\pi}}\, e^{ikx} \qquad k \text{ any real number} \tag{3.145}$$

$$\langle k\,|k'\,\rangle = \delta(k-k')$$

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

We note two corollaries of our proof of $|A| = \sqrt{1/2\pi}$:

$$\int_{-\infty}^{\infty} dx\, e^{ikx} = 2\,\pi\delta(k) \qquad \int_{-\infty}^{\infty} dk\, e^{ikx} = 2\,\pi\delta(x) \tag{3.146}$$

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

Expansion of Vector Space Elements in the K Eigenbasis

We required completeness of our $\{|k\rangle\}$ basis in both the finite and infinite interval cases, so we have

$$\sum_{j=-\infty}^{\infty} |k_j\rangle\langle k_j| = I \qquad\qquad \int_{-\infty}^{\infty} dk\, |k\rangle\langle k| = I$$

We drop the $_L$ and $_\infty$ on I; it should be clear from context which we are discussing. With the above, we may expand any ket in terms of the $\{|k\rangle\}$ basis. We have

$$|f\rangle = \sum_{j=-\infty}^{\infty} |k_j\rangle\langle k_j|f\rangle \qquad\qquad |f\rangle = \int_{-\infty}^{\infty} dk\, |k\rangle\langle k|f\rangle$$

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

The immediate question is: what is $\langle k_j | f \rangle$ or $\langle k | f \rangle$? This is straightforward to calculate using what we know about the $\{|x\rangle\}$ basis and about $\langle x | k \rangle$ (the latter is basically the elements of the unitary matrix that transforms from one basis to the other):

$$\langle k_j | f \rangle = \int_a^b dx \, \langle k_j | x \rangle \langle x | f \rangle = \sqrt{\frac{1}{L}} \int_a^b dx \, e^{-ik_j x} f(x)$$

$$\langle k | f \rangle = \int_{-\infty}^{\infty} dx \, \langle k | x \rangle \langle x | f \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx \, e^{-ikx} f(x)$$

We thus begin to understand why $|f\rangle$ and $f(x)$ are not quite the same thing. One can expand $|f\rangle$ in terms of different bases (*i.e.*, write down different matrix representation): the default basis is the $\{|x\rangle\}$ basis, and the coefficients of the expansion in this basis are (the matrix representation is) $\langle x | f \rangle = f(x)$; but one can also expand $|f\rangle$ in the $\{|k\rangle\}$ basis, and the coefficients of that expansion are (the matrix representation is) $\langle k | f \rangle$.

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

We thus will find it necessary to subscript our "functions" to indicate what basis they assume (what matrix representation they refer to); that is, we write

$$f_x(x) = \langle x \,|f\, \rangle \qquad f_k(k) = \langle k \,|f\, \rangle \tag{3.147}$$

The use of the x or k in both the argument and the subscript may seem redundant, but it allows us to put anything in the argument without ambiguity arising as to which representation for $|f\,\rangle$ we are working with.

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

X and K in the K Eigenbasis

We have defined X and K in the X eigenbasis; their matrix elements are

$$\langle x \,|\, X | x' \rangle = x\,\delta(x - x')$$
$$\langle x \,|\, K \,| x' \rangle = -i \left[\frac{d}{dx} \delta(x - x') \right] = -i\,\delta(x - x') \frac{d}{dx'}$$

where the [ ]'s indicate that the derivative acts only on the δ function.

It is obvious that the matrix elements of K in the K eigenbasis are

$$\langle k \,|\, K \,| k' \rangle = k\,\delta(k - k') \tag{3.148}$$

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

So, how does X act in the K eigenbasis? Let's just calculate it by, as usual, inserting a completeness relation:

$$\begin{aligned}\langle k\,|\,X\,|\,k'\,\rangle &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx'\,\langle k\,|\,x\,\rangle\langle x\,|\,X\,|\,x'\,\rangle\langle x'\,|\,k'\,\rangle \qquad (3.149)\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx'\,e^{-ikx}\,x\,\delta(x-x')e^{ik'x'}\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\,x\,e^{-i(k-k')x} = i\frac{d}{dk}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} dx\,e^{-i(k-k')x}\right]\\ &= i\frac{d}{dk}\delta(k-k') = i\,\delta(k-k')\frac{d}{dk'}\end{aligned}$$

Hence, the action of X on a ket $|f\,\rangle$ is

$$\langle k\,|\,X\,|f\,\rangle = i\,\frac{df_k(k)}{dk} \quad \text{or, somewhat misleadingly,} \quad X|f\,\rangle = \left|\,i\,\frac{df_k(k)}{dk}\,\right\rangle \qquad (3.150)$$

(The latter is misleading because we are trying to divorce the kets from their coordinate (functional) representation in either the $\{|x\,\rangle\}$ or $\{|k\,\rangle\}$ basis.)

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

Finally, let us calculate the interesting operator $[X, K]$:

$$\begin{aligned}
\langle x \,|\, X\,K \,|f\,\rangle &= \int_{-\infty}^{\infty} dx'\langle x \,|\, X|x'\,\rangle\langle x' \,|\, K \,|f\,\rangle = \int_{-\infty}^{\infty} dx'\, x\,\delta(x-x')\,(-i)\,\frac{df}{dx'} \\
&= -i\,x\,\frac{df}{dx} \\
\langle x \,|\, K\,X \,|f\,\rangle &= \int_{-\infty}^{\infty} dx'\langle x \,|\, K|x'\,\rangle\langle x' \,|\, X \,|f\,\rangle = \int_{-\infty}^{\infty} dx'(-i)\,\delta(x-x')\frac{d}{dx'}x'\,f(x') \\
&= -i\int_{-\infty}^{\infty} dx'\,\delta(x-x')\left[f(x') + x'\frac{df(x')}{dx'}\right] = -i\,f(x) - i\,x\frac{df}{dx} \\
&\neq \langle x \,|\, X\,K \,|f\,\rangle
\end{aligned}$$

$$\Longrightarrow \quad \langle x \,|\, [X,K] \,|f\,\rangle = i\,f(x) = i\,\langle x \,|f\,\rangle \quad \Longleftrightarrow \quad [X,K] = i\,I \tag{3.151}$$

Section 4
Postulates Revisited

Lecture 13:
Postulates of Quantum Mechanics Revisited

Date Given: 2008/10/29
Date Revised: 2008/10/29

Summary

Recall the Postulates we briefly discussed in Section 1.2:

1. The state of a particle is represented by a vector in a physical Hilbert space.
2. The fundamental state variables x and p of classical mechanics are replaced by Hermitian operators X and P whose matrix elements are well specified in a physical Hilbert space basis consisting of position eigenstates (states with perfectly defined position x). Any derived dynamical variables $\omega(x, p)$ are replaced by operators Ω defined by the above correspondence.
3. Measurement of any classical variable $\omega(x, p)$ for a quantum state yields only the eigenvalues of the corresponding operator Ω, with the probability of obtaining the eigenvalue ω given by the squared norm of the projection of the state onto the eigenstate corresponding to ω.
4. The state vector evolves according to the Schrödinger equation.

We now have the language to interpret what is meant by these postulates. We do that in this section.

Postulate 1: Representation of Particle States

The state of a particle is represented by a vector $|\psi(t)\rangle$ in a physical Hilbert space.

We now know what is meant by this statement mathematically, in a generic sense: the state $|\psi(t)\rangle$ is an element in an inner product space; members of such spaces have the following important properties:

- They can be added together linearly, with coefficients that are just numbers.
- An inner product is defined that provides for definitions of orthogonality and normalization.
- There exist orthonormal sets of basis states and all states can be written as linear combinations of them.
- In terms of any particular basis, the vector corresponding to the state has a column matrix representation, and the corresponding dual vector in the dual vector space has a row matrix representation that is the conjugate transpose of the column matrix representation. When the basis is orthonormal, the inner product is equivalent to matrix multiplication of the column and row matrices.

Postulate 1: Representation of Particle States (cont.)

- Operators can act on the states and return new states, and any operator has a matrix representation for any particular choice of orthonormal basis.
- There are Hermitian operators that have real eigenvalues and a set of eigenvectors yields as an orthonormal basis.
- There are unitary operators that can be used to rotate from pne orthonormal basis to another and which themselves have unit-modulus eigenvalues and orthonormal sets of eigenvectors.
- The space has been restricted so that the K operator is Hermitian.

Normalization Considerations

One implication of this postulate is that, when we take a linear combination of states, $|\chi\rangle = \alpha|\psi\rangle + \beta|\phi\rangle$, we will in general want to normalize the result; that is, we should define

$$|\chi\rangle = \frac{\alpha|\psi\rangle + \beta|\phi\rangle}{\sqrt{|\alpha|^2 + |\beta|^2}}$$

so that $|\chi|^2 = \langle\chi|\chi\rangle = 1$ if $|\psi|^2 = \langle\psi|\psi\rangle = 1$ and $|\phi|^2 = \langle\phi|\phi\rangle = 1$ (you can check this by writing out $\langle\chi|\chi\rangle$). As we will see, this convention ensures measurement probabilities will be automatically normalized for $|\chi\rangle$.

Postulate 2: Correspondence for Classical Variables

The independent variables x and p that describe completely the state of a particle in classical mechanics are represented by Hermitian operators X and P in the physical Hilbert space of states, with X and P having the following matrix elements when using the position basis for the physical Hilbert space:

$$\langle x\,|X\,|x'\,\rangle = x\delta\left(x-x'\right) \qquad \langle x\,|P\,|x'\,\rangle = -i\,\hbar\,\frac{d}{dx}\delta\left(x-x'\right) \tag{4.1}$$

Any arbitrary classical dynamical variable $\omega(x,p)$ has a corresponding Hermitian operator

$$\Omega(X,P) = \omega(x\to X, p\to P) \tag{4.2}$$

where we simply replace x and p in ω with X and P to obtain $\Omega(X,P)$.

Having been through the exercise of constructing the infinite-dimensional generalization of inner product spaces, we now understand what is meant by the operators X and P and their matrix elements. Postulate 3 tells us how the above matrix elements are related to measurements. The extension to arbitrary classical variables $\omega(x,p)$ is also clear, modulo the issue of having to deal with ambiguous combinations of x and p (*i.e.*, if one has the classical quantity $x\,p$, should one use $X\,P$, $P\,X$, their sum, or their difference?).

Postulate 2: Correspondence for Classical Variables (cont.)

One thing that will not be clear yet, and cannot be discussed until you have seen Hamiltonian mechanics in Ph106a, is why we make the above choice for P. This choice was a clever guess by the creators of quantum mechanics based on analogy to classical mechanics. An alternative version of this postulate makes this more clear: it takes X as defined above but then makes the requirement $[X, P] = i\hbar$. This latter relation is the quantum mechanical analogue of the classical Poisson bracket $\{x, p\} = 1$. Written this way, the path from classical mechanics to quantum mechanics is explicit. But that does not amount to a proof. By definition, a postulate can never be explicitly proven, but it can be motivated and then checked that it gives physically reasonable and correct results in particular situations.

Postulate 3: Results of Measurements of Classical Variables

Let $\{|\omega\rangle\}$ denote the set of eigenstates of the Hermitian operator with eigenvalues ω. If a particle is in an arbitrary state $|\psi\rangle$, then measurement of the variable corresponding to the operator Ω will yield only the eigenvalues $\{\omega\}$ of Ω. The measurement will yield the particular value ω for that variable with relative probability $P(\omega) = |\langle\omega|\psi\rangle|^2$ and the system will change from state $|\psi\rangle$ to state $|\omega\rangle$ as a result of the measurement being made.

Let's break the statement down carefully:

1. The eigenvalues of Ω are the only values the measured quantity may take on.
2. The measurement outcome is fundamentally probabilistic, and the relative probability of a particular allowed outcome ω is given by finding the projection of $|\psi\rangle$ onto the corresponding eigenstate $|\omega\rangle$. By *relative* probability, we simply mean that the ratio of the probabilities of two outcomes is given by $P(\omega_1)/P(\omega_2) = |\langle\omega_1|\psi\rangle|^2 / |\langle\omega_2|\psi\rangle|^2$. The *absolute* probability of a particular outcome requires a normalizing factor that sums over all possible measurement outcomes, to be discussed later. This implies that, if $|\psi\rangle$ is an eigenstate of Ω, then the measurement will always yield the corresponding eigenvalue.
3. The measurement process itself changes the state of the particle to the eigenstate $|\omega\rangle$ corresponding to the measurement outcome ω. This is the equivalent of applying the projection operator P_ω (but one only knows which P_ω to use after the measurement has been done!) and then renormalizing the state.

Postulate 3: Results of Measurements of Classical Variables (cont.)

Where the math ends and the physics starts

As we noted in Section 1.2, we could have made a more classical interpretation of the expansion of $|\psi\rangle$ in terms of the eigenvectors $\{|\omega\rangle\}$: that the result of the measurement would be the weighted sum of the eigenvalues, weighted by the norms of the expansion coefficients $|\langle\omega|\psi\rangle|^2$ rather than $\langle\omega|\psi\rangle$ because the former is guaranteed to be real while the latter is not. But we do not do that. It is a physical assumption that the expansion coefficients are to be interpreted as probabilities of the allowed outcomes, not as weighting factors.

Also, we could have assumed that measurement is not an operation that changes $|\psi\rangle$: we could have said that $|\psi\rangle$ evolves in some way independent of any measurements that take place. Even if we had said that the action of a measurement on $|\psi\rangle$ is to act with the corresponding operator Ω on $|\psi\rangle$, we would not arrive at this postulate. It is entirely **outside of the mathematical structure** to assume that a measurement to which the operator Ω corresponds results in $|\psi\rangle$ collapsing to one of the $\{|\omega\rangle\}$ via application of the appropriate P_ω. It is an assumption that yields correct predictions for experimental results.

Note that there is no way to write the collapse process explicitly as an operator. The problem is that you only know which P_ω to apply after you have obtained the measured value ω; the act of measurement does not, for example, correspond to the operator $M = \sum_j P(\omega_j)P_{\omega_j} = \sum_j |\langle\omega_j|\psi(t)\rangle|^2|\omega_j\rangle\langle\omega_j|$.

Postulate 3: Results of Measurements of Classical Variables (cont.)

Degeneracy

In the case of degenerate eigenvalues, the obvious generalization of the above postulate is to replace $P(\omega) = |\langle \omega \,|\psi\rangle|^2$ with

$$P(\omega) = |P_\omega|\psi\rangle|^2 = \langle \psi \,|\, P_\omega \, P_\omega|\psi\rangle = \langle \psi \,|\, P_\omega|\psi\rangle \tag{4.3}$$

where P_ω is the projection operator for the ω subspace,

$$P_\omega = \sum_{\omega_j=\omega} |\omega_j\rangle\langle\omega_j| \tag{4.4}$$

and where we have written out three equivalent expressions using the fact that the projection operator P_ω is Hermitian and satisfies $P_\omega^2 = P_\omega$. This expression results in

$$P(\omega) = \sum_{\omega_j=\omega} |\langle \omega_j \,|\psi\rangle|^2 \tag{4.5}$$

That is, when one has degenerate eigenvalues, the relative probability of obtaining a degenerate eigenvalue is the sum of the relative probabilities for all the states corresponding to that eigenvalue.

Postulate 3: Results of Measurements of Classical Variables (cont.)

In the absence of degeneracy, the above generalization simplifies to our original postulate because $P_\omega = |\omega\rangle\langle\omega|$:

$$|\langle\omega|\psi\rangle|^2 = \langle\psi|\omega\rangle\langle\omega|\psi\rangle = \langle\psi|P_\omega|\psi\rangle = \langle\psi|P_\omega P_\omega|\psi\rangle = |P_\omega|\psi\rangle|^2 \tag{4.6}$$

An interesting thing about degeneracy is that states may not completely collapse upon measurement. The measurement will apply the projection operator from Equation 4.4. Therefore, if the particle begins in a state that has non-zero expansion coefficients for more than one of the $|\omega_j\rangle$, then it will retain those nonzero expansion coefficients for all $|\omega_j\rangle$ that correspond to the measured eigenvalue ω. That is,

$$\text{if} \qquad |\psi\rangle = \sum_{\omega_j=\omega} c_j|\omega_j\rangle + \sum_{\omega_j\neq\omega} c_j|\omega_j\rangle$$

$$\text{then} \quad P_\omega|\psi\rangle = \left[\sum_{\omega_k=\omega} |\omega_k\rangle\langle\omega_k|\right]\left[\sum_{\omega_j=\omega} c_j|\omega_j\rangle + \sum_{\omega_j\neq\omega} c_j|\omega_j\rangle\right] = \sum_{\omega_j=\omega} c_j|\omega_j\rangle$$

The state is collapsed to the subspace $\mathbb{V}_\omega$, but not to a single eigenstate $|\omega\rangle$. One then has to renormalize the resulting state.

Postulate 3: Results of Measurements of Classical Variables (cont.)

Normalization of probabilities

Let us consider three cases:

- **finite-dimensional case:**
  For the finite-dimensional case, the assumption that the relative probability of outcome ω is given by $|\langle \omega | \psi \rangle|^2$, combined with the very reasonable assumption that there must be *some* outcome, immediately implies that that absolute probability of outcome ω is

$$P(\omega_j) = \frac{|\langle \omega_j | \psi \rangle|^2}{\sum_{j=1}^{n} |\langle \omega_j | \psi \rangle|^2} \tag{4.7}$$

  In fact, for a properly normalized state, the denominator is trivial:

$$\sum_{j=1}^{n} |\langle \omega_j | \psi \rangle|^2 = \sum_{j=1}^{n} \langle \psi | \omega_j \rangle \langle \omega_j | \psi \rangle = \langle \psi | \psi \rangle = 1$$

  via the completeness relation $I = \sum_{j=1}^{n} |\omega_j \rangle \langle \omega_j|$.

Postulate 3: Results of Measurements of Classical Variables (cont.)

- **infinite-dimensional, but considering an operator whose eigenvalues are discretized (though possibly infinite in number)**
  e.g., the K operator for our example of functions on the interval $[a, b]$: the above rule continues to hold exactly. The denominator is guaranteed to remain finite in spite of the infinite sum because it is the normalization of the state $|\psi\rangle$.
- **infinite-dimensional case and considering an operator whose eigenvalues are a continuum (uncountably infinite)**
  e.g., the X operator for our example of functions on the interval $[a, b]$: we must reinterpet the expansion coefficients as a probability density. That is, the probability of obtaining from the measurement corresponding to Ω a value between ω and $\omega + d\omega$ is

$$P(\omega)\, d\omega = \frac{|\langle \omega | \psi \rangle|^2 d\omega}{\int_{\omega_-}^{\omega_+} d\omega\, |\langle \omega | \psi \rangle|^2} \tag{4.8}$$

where ω_- and ω_+ are the minimum and maximum allowed values of ω, which might be $\pm\infty$. Equivalently, the probability of obtaining a value in the interval $[\omega_1, \omega_2]$ is

$$P(\omega_1 \leq \omega \leq \omega_2) = \frac{\int_{\omega_1}^{\omega_2} d\omega\, |\langle \omega | \psi \rangle|^2}{\int_{\omega_-}^{\omega_+} d\omega\, |\langle \omega | \psi \rangle|^2} \tag{4.9}$$

Postulate 3: Results of Measurements of Classical Variables (cont.)

What happens when we have states that are delta-function normalized? For our example $\Omega = X$, and $|\psi\rangle = |x_0\rangle$ and eigenstate, we have

$$\int_a^b dx\, |\langle x|\psi\rangle|^2 = \int_a^b dx\, \langle x_0|x\rangle\langle x|x_0\rangle = \int_a^b dx\, \delta(x_0 - x)\,\delta(x - x_0) = \delta(0)$$

which is infinite. Let's reconsider this expression in the context of Equation 4.9, moving the normalizing factor to the left side:

$$\left[\int_a^b dx\, |\langle x|x_0\rangle|^2\right] P(x_1 \le x \le x_2) = \int_{x_1}^{x_2} dx\, |\langle x|x_0\rangle|^2$$

$$\delta(0)P(x_1 < x < x_2) = \begin{cases} \delta(0) & \text{if } x_1 \le x_0 \le x_2 \\ 0 & \text{if } x_0 < x_1 \text{ or } x_0 > x_2 \end{cases}$$

$$P(x_1 < x < x_2) = \begin{cases} 1 & \text{if } x_1 \le x_0 \le x_2 \\ 0 & \text{if } x_0 < x_1 \text{ or } x_0 > x_2 \end{cases}$$

where we have taken the somewhat unrigorous step of dividing both sides by $\delta(0)$ (both sides could be put under an integral sign with an arbitrary function to be more rigorous). The point is that, while the normalizing factor is formally infinite, this is only a problem when one considers the differential expression; one can obtain reasonable results for the probability in any finite interval, which is what is experimentally accessible.

Postulate 3: Results of Measurements of Classical Variables (cont.)

The above analysis works even if $x_1 \to -\infty$ and $x_2 \to +\infty$.

One runs into a similar problem with the K eigenstates for the infinite interval, but again one can obtain sensible results by only considering probabilities integrated over some finite range.

Moreover, **relative** probabilities are always well-defined because the infinite normalizing denominator cancels out. For our example, one still gets infinities, but they are sensible infinities: if, given a particle in the eigenstate $|x_0\rangle$, one wants to compare the probability of the particle being in the intervals $[x_1, x_2]$ and $[x_3, x_4]$ that do not overlap, clearly it can only be in one or the other, so the ratio of the two probabilities must either be infinite or zero.

In addition, we can always create states that are reasonably normalized – these simply will not be eigenstates of X or P.

Wavefunctions

Given some continuous eigenvalue ω, the quantity $\langle \omega | \psi \rangle$ can be considered a function of the continuous variable ω. It is conventional to call this quantity the wavefunction and write it as $\psi(\omega)$. The most common use of this nomenclature is for $\psi(x) = \langle x | \psi \rangle$, but it could also be used for $\psi_k(k) = \langle k | \psi \rangle$ when k is continous. The use of a notation like $\psi(\omega)$ can be confusing because the function ψ is different depending on which operator the eigenvalues correspond to — *e.g.*, above, $\psi(x)$ and $\psi_k(k)$ are in general very different functions — so, the argument of the function, which is normally a dummy variable, means something. To be clear, we will frequently use the labeling $\psi_\omega(\omega) = \langle \omega | \psi \rangle$ so that $\psi_x(x) = \langle x | \psi \rangle$ and $\psi_k(k) = \langle k | \psi \rangle$. This makes it clear what basis we are projecting onto and thus what quantity the "wavefunction" is a function of.

We will refer to $\psi_x(x) = \langle x | \psi \rangle$ as the *position-space* or *coordinate-space wavefunction* and $\psi_k(k) = \langle k | \psi \rangle$ as the *k-space*, or once we have added a $\hbar$ to turn K into the momentum operator P, the *momentum-space wavefunction*. When the eigenvalue is discretized, such as k on a finite interval $[a, b]$, we tend not to use the "wavefunction" language, but this is just semantics and a reluctance to call a quantity that is defined on discrete points a function. There is no truly fundamental difference between a "wavefunction" and the set of expansion coefficients of a state in a particular basis.

Commuting and Non-Commuting Operators

We now also see the physical relevance of whether two operators corresponding to physical observables commute. Let us first neglect degeneracies. If two Hermitian operators Ω and Λ commute, then, as we proved in Section 3.6, there is a set of common eigenstates $\{|j\rangle\}$ that have eigenvalues $\{\omega_j\}$ and $\{\lambda_j\}$. If $|\psi\rangle$ is an eigenstate $|j\rangle$, then measurements of Ω and Λ will yield the definite values ω_j and λ_j. If $|\psi\rangle$ is not an eigenstate, then the measurement outcomes will be correlated: if Ω yields ω_j, then Λ yields λ_j because the projection operator $P_{\omega=\omega_j}$ is the same as the projection operator $P_{\lambda=\lambda_j}$. The relative probabilities $P(\omega_j)$ and $P(\lambda_j)$ will of course be equal.

If there are degeneracies, then the correspondence may break down because of incompletely overlapping subspaces. But this is completely consistent with the above statement; what occurs would just be the result of there being multiple eigenstates that contribute to a given $P(\omega)$.

Our archetypal example of two non-commuting operators is X and P, which we proved in Section 3.9 (up to a factor of $\hbar$) gives $[X, P] = i\hbar$. These clearly do not commute, implying that there are no states that have definite values of both X and P.

Expectation Values and Uncertainties

Because measurement outcomes are probabilistic, the next most definite quantities to consider are probability-weighted moments of the measurements. The expectation value of an operator Ω is simply the probability-weighted mean outcome,

$$\langle \Omega \rangle = \sum_j P(\omega_j)\,\omega_j \quad \text{or} \quad \langle \Omega \rangle = \int_{\omega_-}^{\omega_+} d\omega\, P(\omega)\,\omega \tag{4.10}$$

We can write this explicitly in terms of Ω and the state $|\psi\rangle$:

$$\begin{aligned} \langle \Omega \rangle &= \sum_j |\langle \omega_j | \psi \rangle|^2 \omega_j = \sum_j \langle \psi | \omega_j \rangle \langle \omega_j | \psi \rangle \omega_j = \sum_j \langle \psi | \Omega | \omega_j \rangle \langle \omega_j | \psi \rangle \\ &= \langle \psi | \Omega | \psi \rangle \end{aligned} \tag{4.11}$$

where we used completeness to make the last step. A similar derivation holds for the continuous ω version so that the same result $\langle \Omega \rangle = \langle \psi | \Omega | \psi \rangle$ holds.

Postulate 3: Results of Measurements of Classical Variables (cont.)

The next moment to consider is the variance of ω, which is conventionally defined as

$$\langle(\Delta\Omega)^2\rangle = \sum_j P(\omega_j)\left(\omega_j - \langle\Omega\rangle\right)^2 \quad \text{or} \quad \langle(\Delta\Omega)^2\rangle = \int_{\omega_-}^{\omega_+} d\omega\, P(\omega)\left(\omega - \langle\Omega\rangle\right)^2 \tag{4.12}$$

Let's pursue this in the discretized case. First, the above expression can be simplified:

$$\begin{aligned}\sum_j P(\omega_j)\left(\omega_j - \langle\Omega\rangle\right)^2 &= \sum_j P(\omega_j)\omega_j^2 - 2\langle\Omega\rangle\sum_j P(\omega_j)\omega_j + \langle\Omega\rangle^2\sum_j P(\omega_j) \\ &= \left[\sum_j P(\omega_j)\,\omega_j^2\right] - \langle\Omega\rangle^2\end{aligned}$$

where we used the definition of $\langle\Omega\rangle$ to reduce the second term and the normalization of the probability to reduce the third term.

Postulate 3: Results of Measurements of Classical Variables (cont.)

Let's write out the first term in terms of Ω and $|\psi\rangle$.

$$\sum_j P(\omega_j)\,\omega_j^2 = \sum_j \langle\psi|\omega_j\rangle\langle\omega_j|\psi\rangle\,\omega_j^2 = \sum_j \langle\psi|\Omega|\omega_j\rangle\langle\omega_j|\Omega|\psi\rangle = \langle\psi|\Omega^2|\psi\rangle$$

where we again used completeness. So we have

$$\langle(\Delta\Omega)^2\rangle = \langle\psi|\Omega^2|\psi\rangle - \langle\Omega\rangle^2 = \langle\psi|\left[\Omega^2 - \langle\Omega\rangle^2\right]|\psi\rangle = \langle\psi|\left[\Omega - \langle\Omega\rangle\right]^2|\psi\rangle \quad (4.13)$$

where we have written three algebraically equivalent forms by using $\langle\psi|\psi\rangle = 1$ and the kind of conversion between $\Omega^2 - \langle\Omega\rangle^2$ and $[\Omega - \langle\Omega\rangle]^2$ that we used on the previous page.

Lecture 14:
Postulates of Quantum Mechanics Revisited, Continued
One-Dimensional Free Particle Propagator

Date Given: 2008/10/31
Date Revised: 2008/10/31

Postulate 4: Time Evolution of States

The time evolution of the state vector $|\psi(t)\rangle$ is governed by the Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H|\psi(t)\rangle \tag{4.14}$$

where $H(X, P)$ is the operator obtained from the classical Hamiltonian $\mathcal{H}(x, p)$ via the correspondence $x \rightarrow X$ and $p \rightarrow P$.

For most systems we will consider H, will be the energy of the system and will be independent of time. We will consider some more complicated cases, and we will revisit the distinction between $\mathcal{H}$ and the energy when we encounter them.

The above equation for the time evolution of $|\psi\rangle$ allows us to write a fairly definite form for $|\psi(t)\rangle$. There are two versions:

Postulate 4: Time Evolution of States (cont.)

- **Generic operator form**
  Let us assume there is an operator $U(t)$ (recall this is called the propagator) such that $|\psi(t)\rangle = U(t)|\psi(0)\rangle$. Then the above equation becomes

$$i\hbar \frac{dU(t)}{dt} |\psi(0)\rangle = H\, U(t)\, |\psi(0)\rangle$$

  $|\psi(0)\rangle$ is constant and arbitrary, so we obtain a differential equation for $U(t)$:

$$i\hbar \frac{dU(t)}{dt} = H\, U(t) \tag{4.15}$$

  Our discussion of calculus with operators (Section 9) tells us the solution is

$$U(t) = e^{-\frac{i}{\hbar} H\, t}$$

  Thus, our full solution for the time evolution of the state is

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle = e^{-\frac{i}{\hbar} H\, t}|\psi(0)\rangle \tag{4.16}$$

  Since we assume H is Hermitian, we are guaranteed the propagator $U(t)$ is unitary and thus preserves the norm $|\psi|^2$.

Postulate 4: Time Evolution of States (cont.)

- **Eigenbasis form**
  H is a Hermitian operator, so it has eigenvalues $\{E_j\}$ (discrete or continuous). If we consider an eigenstate $|E_j\rangle$, then we have

$$H|E_j\rangle = E_j|E_j\rangle$$

for all time (assuming time-independent H). Thus, for this state, our differential equation becomes

$$i\hbar\frac{d}{dt}|E_j(t)\rangle = H|E_j(t)\rangle = E_j|E_j(t)\rangle$$

Since $|E_j(t)\rangle$ has to remain an eigenvector of H, only its coefficient may change. The differential equation suggests the solution $|E_j(t)\rangle = e^{-\frac{i}{\hbar}E_j t}|E_j(0)\rangle$; one can check trivially that it satisfies the equation.

Postulate 4: Time Evolution of States (cont.)

With this in hand, we can calculate the time evolution of any state:

$$|\psi(t)\rangle = \sum_j |E_j(t)\rangle\langle E_j(t)|\psi(t)\rangle$$

This does not get us very far yet, but let's note something useful:

$$\begin{aligned}
\frac{d}{dt}\langle E_j(t)|\psi(t)\rangle &= \left[\frac{d}{dt}\langle E_j(t)|\right]|\psi(t)\rangle + \langle E_j(t)|\left[\frac{d}{dt}|\psi(t)\rangle\right] \\
&= \left[\frac{d}{dt}|E_j(t)\rangle\right]^\dagger|\psi(t)\rangle + \langle E_j(t)|\left[-\frac{i}{\hbar}H|\psi(t)\rangle\right] \\
&= \left[-\frac{i}{\hbar}H|E_j(t)\rangle\right]^\dagger|\psi(t)\rangle - \frac{i}{\hbar}\langle E_j(t)|H|\psi(t)\rangle \\
&= \left[-\frac{i}{\hbar}E_j|E_j(t)\rangle\right]^\dagger|\psi(t)\rangle - \frac{i}{\hbar}\left[H|E_j(t)\rangle\right]^\dagger|\psi(t)\rangle \\
&= \frac{i}{\hbar}E_j\langle E_j(t)|\psi(t)\rangle - \frac{i}{\hbar}\left[E_j|E_j(t)\rangle\right]^\dagger|\psi(t)\rangle \\
&= \frac{i}{\hbar}E_j\langle E_j(t)|\psi(t)\rangle - \frac{i}{\hbar}E_j\langle E_j(t)|\psi(t)\rangle = 0
\end{aligned}$$

Postulate 4: Time Evolution of States (cont.)

We see that the projection of $|\psi(t)\rangle$ onto $|E_j(t)\rangle$ is time-independent. We could have seen this from the fact that the same unitary operator $U(t)$ time-evolves both of them and thus their inner-product is time independent, but we did not want to assume our other proof in doing this proof. With this fact, we may conclude $\langle E_j(t)|\psi(t)\rangle = \langle E_j(0)|\psi(0)\rangle$ and thus

$$|\psi(t)\rangle = \sum_j |E_j(t)\rangle\langle E_j(0)|\psi(0)\rangle = \sum_j e^{-\frac{i}{\hbar}E_j t}|E_j(0)\rangle\langle E_j(0)|\psi(0)\rangle \quad (4.17)$$

Combining the two forms by noting that $|\psi(0)\rangle$ is arbitrary, or also by using the bilinear form for U, we have an explicit form for the propagator $U(t)$:

$$U(t) = e^{-\frac{i}{\hbar}Ht} = \sum_j e^{-\frac{i}{\hbar}E_j t}|E_j(0)\rangle\langle E_j(0)| = \sum_j e^{-\frac{i}{\hbar}E_j t}P_{E_j(0)} \quad (4.18)$$

where $P_{E_j(0)}$ is the projection operator onto the eigenvector $|E_j(0)\rangle$. The is similar to the propagator we found for Example 6, though here we do not need to take the real part, and, because the Schrödinger Equation is first-order, $|\dot{\psi}(0)\rangle$ is not involved.

We will in general drop the (0) in $|E_j(0)\rangle$ and $\langle E_j(0)|$ when H is time-independent as we have assumed here.

Postulate 4: Time Evolution of States (cont.)

For a time-dependent Hamiltonian $H(t)$, such a simple form does not hold because our differential equation for $U(t)$, Equation 4.15, now becomes

$$i\,\hbar\,\frac{dU(t)}{dt} = H(t)\,U(t)$$

The solution is no longer $U(t) = e^{-\frac{i}{\hbar}\,H\,t}$: if one takes the time derivative, one gets

$$i\,\hbar\,\frac{dU(t)}{dt} = \left[H(t) + t\,\frac{dH(t)}{dt}\right]U(t)$$

Rather, as Shankar shows, the solution becomes

$$U(t) = \mathcal{T}\left[\exp\left(-\frac{i}{\hbar}\int_0^t dt'\,H(t')\right)\right] = \lim_{N\to\infty}\prod_{j=0}^{N-1}\exp\left[-\frac{i}{\hbar}\left(\frac{t}{N}\right)H\left(\frac{j}{N}\,t\right)\right] \tag{4.19}$$

where $\mathcal{T}[\ ]$ denotes the time-ordered integral, the infinite product of time evolution over infinitesimally small intervals for which the time-independent solution holds. Since we will not consider time-dependent problems in ph125ab, we will not derive the above.

Section 5
Simple One-Dimensional Problems

The Free-Particle Hamiltonian and its Eigenvalues and Eigenvectors

Classical mechanics tells us that the Hamiltonian function for the free particle in one dimension is $\mathcal{H}(x,p) = \frac{p^2}{2\,m}$. Thus, our quantum mechanical Hamiltonian operator is

$$H = \frac{P^2}{2\,m} \tag{5.1}$$

Our next step is to find the eigenvalues and eigenvectors of H; as explained in Section 4.5, once we know the time evolution of the eigenvectors, we can decompose any initial state in the eigenbasis of H, time evolve each eigenvector, and then reconstruct the full time-evolved state. This entire procedure is summarized in Equation 4.18, reproduced here:

$$|\psi(t)\,\rangle = U(t)|\psi(0)\,\rangle \qquad U(t) = e^{-\frac{i}{\hbar}\,H\,t} = \sum_i e^{-\frac{i}{\hbar}\,E_i\,t}|E_i(0)\,\rangle\langle E_i(0)\,|$$

The Free Particle (cont.)

In this case, our work is reduced because we already know what the eigenbasis of H is: it is the same as the eigenbasis of K because $P = \hbar K$ and $H = \frac{P^2}{2m}$. We shall relabel this basis $\{|p\rangle\}$ with the eigenvalue correspondence $p = \hbar k$ simply to avoid confusion; but we emphasize that $|p\rangle \propto |k\rangle$ if $p = \hbar k$: the two vectors are in the same direction in the Hilbert space. (Actually, $|p\rangle = \frac{1}{\sqrt{\hbar}}|k\rangle$ so that $\langle p|p'\rangle = \delta(p - p')$ is consistent with $\langle k|k'\rangle = \delta(k - k')$: the two delta functions differ by a factor of $\hbar$ due to the properties of the delta function.) Because the eigenvalues k of K may be any real number, the eigenvalues p of P may be any real number. The eigenvalues of H are given by acting on these eigenstates with H:

$$H|p\rangle = \frac{P^2}{2m}|p\rangle = \frac{p^2}{2m}|p\rangle \equiv E|p\rangle \tag{5.2}$$

Since p may be any real number, E may be any nonnegative real number. We see that there is twofold degeneracy: one gets the same E for $|+p\rangle$ and $|-p\rangle$. To be clear, we adopt the labeling

$$|E_+\rangle = |p = +\sqrt{2mE}\rangle \quad |E_-\rangle = |p = -\sqrt{2mE}\rangle \tag{5.3}$$

The Free-Particle Propagator

With the eigenvalues and eigenvectors of H in hand, we can calculate the propagator:

$$U(t) = \int_{-\infty}^{\infty} dp\, e^{-\frac{i}{\hbar}\,\frac{p^2}{2\,m}\,t} |p\,\rangle\langle p\,| \tag{5.4}$$

where $|p\,\rangle$ and $\langle p\,|$ we take as shorthand for $|p(t=0)\,\rangle$ and $\langle p(t=0)\,|$ as explained in connection with Equation 4.18.

We note as an aside that the above is **not** the same as

$$U(t) = \int_{0}^{\infty} dE\, e^{-\frac{i}{\hbar}\,E\,t} \left(|E_+\,\rangle\langle E_+\,| + |E_-\,\rangle\langle E_-\,|\right)$$

Even though the kets and bras are in one-to-one correspondence, the integration element dE is not the same as dp; in fact, because $E = p^2/2\,m$, we have $dE = p\,dp/m$. We are certain that the form in terms of p is correct because we have the completeness relation $I = \int_{-\infty}^{\infty} dp\,|p\,\rangle\langle p\,|$. See Shankar Exercise 5.1.1 for more on this point.

The Free Particle (cont.)

It is useful to write down the matrix elements of $U(t)$ in the $\{|x\rangle\}$ representation because that is where we will typically employ it. This is straightforward:

$$\begin{aligned}
[U(t)]_{xx'} = \langle x \,|\, U(t) \,|x'\rangle &= \int_{-\infty}^{\infty} dp\, e^{-\frac{i}{\hbar}\frac{p^2}{2m}t} \langle x\,|p\rangle\langle p\,|x'\rangle \\
&= \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} dp\, e^{-\frac{i}{\hbar}\frac{p^2}{2m}t} e^{\frac{i}{\hbar}p(x-x')} \\
&= \sqrt{\frac{m}{2\pi\hbar i\, t}}\, e^{\frac{i}{\hbar}\frac{m(x-x')^2}{2t}} \qquad (5.5)
\end{aligned}$$

The $(2\pi\hbar)^{-1}$ enters from the normalization of the $|p\rangle$ states, which differ from the normalization of the $|k\rangle$ states by a factor of $\hbar^{-1/2}$. Proving the last step is a bit of nontrivial calculus; one needs to complete the square in the argument of the exponential so that it becomes a perfect quadratic (the counterterm one puts in is independent of p and comes outside the integral as the exponential in $(x-x')^2$) and then one uses the general result $\int_{-\infty}^{\infty} du\, e^{-u^2} = \sqrt{\pi}$, which holds even if u is complex. This is discussed in Shankar Appendix A.2. Also, the complex normalization factor should not be too disturbing; whenever one calculates an observable quantity, one will take a squared modulus, making such factors into real numbers.

Interpretation of the Free-Particle Propagator

First of all, we note that our derivation would apply generally for evolution of $|\psi(t)\rangle$ to $|\psi(t')\rangle$, even for $t' < t$, with the rewriting of U as

$$\left[U(t'-t)\right]_{xx'} = \sqrt{\frac{m}{2\pi\hbar i\,(t'-t)}}\, e^{\frac{i}{\hbar}\frac{(x-x')^2}{2(t'-t)}}$$

Second, we note that we can interpret $[U(t)]_{xx'}$ as the $\{|x\rangle\}$-basis representation of the state (*i.e.*, what we frequently call the *position-basis wavefunction* $\psi_x(x,t)$) one gets at time t if one's initial state is $|x'\rangle$, which has $\{|x\rangle\}$-basis representation $\langle x|x'\rangle = \delta(x-x')$:

$$\psi_x(x,t) = \langle x|\psi(t)\rangle = \langle x|U(t)|\psi(0)\rangle = \langle x|U(t)|x'\rangle = [U(t)]_{xx'}$$

Since we interpret $|\langle x|\psi\rangle|^2$ as the relative probability of the particle's position being in the interval $(x, x+dx)$, it holds that $|\,[U(t)]_{xx'}\,|^2 dx$ is the probability that a particle that is perfectly localized to x' at $t=0$ will be detected in $(x, x+dx)$ at time t.

The Free Particle (cont.)

More generally, U tells us how the probability of finding a particle at position x at time t is determined by the initial wavefunction $\psi_x(x, t = 0) = \langle x\,|\psi(0)\,\rangle$:

$$\begin{aligned}\psi_x(x,t) = \langle x\,|\psi(t)\,\rangle &= \langle x\,|U(t)|\psi(0)\,\rangle \int_{-\infty}^{\infty} dx'\langle x\,|U(t)|x'\,\rangle\langle x'\,|\psi(0)\,\rangle \\ &= \int_{-\infty}^{\infty} dx'\langle x\,|U(t)|x'\,\rangle\psi_x(x,t=0) \\ &= \int_{-\infty}^{\infty} dx'\sqrt{\frac{m}{2\,\pi\,\hbar\,i\,t}}\,e^{\frac{i}{\hbar}\frac{m\,(x-x')^2}{2t}}\psi_x(x',t=0) \qquad (5.6)\end{aligned}$$

The propagator belongs to a class of functions called Green's functions that do similar things: solve some differential equation for a delta-function initial state (or boundary condition) and thereby, by integrating the Green's function with the initial state (boundary condition), one can obtain the full solution at some other time (or position).

The Free Particle (cont.)

We may obtain a similar, but simpler, expression, in the momentum basis because U is diagonal there:

$$[U(t)]_{pp'} = \langle p\,|U(t)|p'\,\rangle = e^{-\frac{i}{\hbar}\frac{p^2}{2m}t}\langle p\,|p'\,\rangle = e^{-\frac{i}{\hbar}\frac{p^2}{2m}t}\delta(p-p') \tag{5.7}$$

Thus, we see if a particle is in an initial state $|p'\,\rangle$ of well-defined momentum, the propagator is trivial: multiply by a complex phase factor. The above is the momentum-space Green's function. Applying it to an arbitrary initial state, we have

$$\begin{aligned}
\psi_p(p,t) &= \langle p\,|\psi(t)\,\rangle = \langle p\,|U(t)|\psi(0)\,\rangle \\
&= \int_{-\infty}^{\infty} dp'\langle p\,|U(t)|p'\,\rangle\langle p'\,|\psi(0)\,\rangle \\
&= \int_{-\infty}^{\infty} dp' e^{-\frac{i}{\hbar}\frac{p^2}{2m}t}\delta(p-p')\psi_p(p,t=0) \\
&= e^{-\frac{i}{\hbar}\frac{p^2}{2m}t}\psi_p(p,t=0)
\end{aligned} \tag{5.8}$$

The integral goes away because each $|p\,\rangle$ eigenbasis element evolves independently.

The Free Particle (cont.)

Finally, let's write a different version of Equation 5.6 to see how the $\{|x\rangle\}$ and $\{|p\rangle\}$ representations of the propagator are related:

$$\begin{aligned}
\psi_x(x,t) &= \langle x|\psi(t)\rangle = \langle x|U(t)|\psi(0)\rangle \\
&= \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dp' \int_{-\infty}^{\infty} dx' \langle x|p\rangle\langle p|U(t)|p'\rangle\langle p'|x'\rangle\langle x'|\psi(0)\rangle \\
&= \int_{-\infty}^{\infty} dx' \langle x'|\psi(0)\rangle \int_{-\infty}^{\infty} dp \langle x|p\rangle \int_{-\infty}^{\infty} dp' e^{-\frac{i}{\hbar}\frac{p^2}{2m}t} \delta(p-p')\langle p'|x'\rangle \\
&= \int_{-\infty}^{\infty} dx' \langle x'|\psi(0)\rangle \int_{-\infty}^{\infty} dp \langle x|p\rangle e^{-\frac{i}{\hbar}\frac{p^2}{2m}t} \langle p|x'\rangle \\
&= \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dx' e^{-\frac{i}{\hbar}\frac{p^2}{2m}t} e^{\frac{i}{\hbar}p(x-x')} \psi_x(x', t=0) \qquad (5.9)
\end{aligned}$$

which ties Equations 5.4, 5.5, 5.6, 5.7, and 5.8 together.

Lecture 15:
One-Dimensional Free Particle: Gaussian Wave Packets

Date Given: 2008/11/03
Date Revised: 2008/11/03

Gaussian Wave Packets

The Gaussian wave packet initial state is one of the few states for which both the $\{|x\rangle\}$ and $\{|p\rangle\}$ basis representations are simple analytic functions and for which the time evolution in either representation can be calculated in closed analytic form. It thus serves as an excellent example to get some intuition about the Schrödinger equation.

We define the $\{|x\rangle\}$ representation of the initial state to be

$$\psi_x(x, t=0) = \langle x|\psi(0)\rangle = \left(\frac{1}{2\pi\sigma_x^2}\right)^{1/4} e^{\frac{i}{\hbar}p_0 x} e^{-\frac{x^2}{4\sigma_x^2}} \tag{5.10}$$

The relation between our σ_x and Shankar's Δ_x is $\Delta_x = \sigma_x\sqrt{2}$. As we shall see, we choose to write in terms of σ_x because $\langle(\Delta X)^2\rangle = \sigma_x^2$.

The Free Particle (cont.)

Before doing the time evolution, let's better understand the initial state. First, the symmetry of $\langle x\,|\psi(0)\,\rangle$ in x implies $\langle X\rangle_{t=0} = 0$, as follows:

$$\begin{aligned}\langle X\rangle_{t=0} = \langle\psi(0)\,|X\,|\psi(0)\,\rangle &= \int_{-\infty}^{\infty} dx\,\langle\psi(0)\,|X\,|x\,\rangle\langle x\,|\psi(0)\,\rangle \\ &= \int_{-\infty}^{\infty} dx\,\langle\psi(0)\,|x\,\rangle\,x\,\langle x\,|\psi(0)\,\rangle \\ &= \int_{-\infty}^{\infty} dx\,x\left(\frac{1}{2\,\pi\,\sigma_x^2}\right)^{1/2} e^{-\frac{x^2}{2\,\sigma_x^2}} = 0 \qquad (5.11)\end{aligned}$$

because the integrand is odd.

Second, we can calculate the initial variance $\langle(\Delta X)^2\rangle_{t=0}$:

$$\langle(\Delta X)^2\rangle_{t=0} = \int_{-\infty}^{\infty} dx\,\left(x^2 - \langle X\rangle_{t=0}^2\right)\left(\frac{1}{2\,\pi\,\sigma_x^2}\right)^{1/2} e^{-\frac{x^2}{2\,\sigma_x^2}} = \sigma_x^2 \qquad (5.12)$$

where we have skipped a few steps that are similar to what we did above for $\langle X\rangle_{t=0}$ and we did the final step using the Gaussian integral formulae from Shankar and the fact that $\langle X\rangle_{t=0} = 0$.

The Free Particle (cont.)

We calculate the $\{|p\rangle\}$-basis representation of $|\psi(0)\rangle$ so that calculation of $\langle P\rangle_{t=0}$ and $\langle(\Delta P)^2\rangle_{t=0}$ are easy (by contrast, Shankar Example 4.2.4 does this in the $\{|x\rangle\}$ basis):

$$\begin{aligned}\psi_p(p,t=0) = \langle p|\psi(0)\rangle &= \int_{-\infty}^{\infty} dx\,\langle p|x\rangle\langle x|\psi(0)\rangle \\ &= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} dx\, e^{-\frac{i}{\hbar}px}\left(\frac{1}{2\pi\sigma_x^2}\right)^{1/4} e^{\frac{i}{\hbar}p_0 x} e^{-\frac{x^2}{4\sigma_x^2}} \\ &= \left(\frac{1}{2\pi\sigma_p^2}\right)^{1/4} e^{-\frac{(p-p_0)^2}{4\sigma_p^2}} \end{aligned} \tag{5.13}$$

where the $\frac{1}{\sqrt{\hbar}}$ comes from the normalization $|p\rangle = \frac{1}{\sqrt{\hbar}}|k\rangle$, where $\sigma_p \equiv \frac{\hbar}{2\sigma_x}$, and the final step is done by completing the square in the argument of the exponential and using the usual Gaussian integral $\int_{-\infty}^{\infty} du\, e^{-u^2} = \sqrt{\pi}$. With the above form for the $\{|p\rangle\}$-space representation of $|\psi(0)\rangle$, the calculation of $\langle P\rangle_{t=0}$ and $\langle(\Delta P)^2\rangle_{t=0}$ are calculationally equivalent to what we already did for $\langle X\rangle_{t=0}$ and $\langle(\Delta X)^2\rangle_{t=0}$, yielding

$$\langle P\rangle_{t=0} = p_0 \qquad \langle(\Delta P)^2\rangle_{t=0} = \sigma_p^2 \tag{5.14}$$

The Free Particle (cont.)

We may now calculate $|\psi(t)\rangle$. Shankar does this only in the $\{|x\rangle\}$ basis, but we do it in the $\{|p\rangle\}$ basis too to illustrate how simple it is in the eigenbasis of H. The result is of course

$$\psi_p(p,t) = \langle p|\psi(t)\rangle = \left(\frac{1}{2\pi\sigma_p^2}\right)^{1/4} e^{-\frac{(p-p_0)^2}{4\sigma_p^2}} e^{-\frac{i}{\hbar}\frac{p^2}{2m}t} \tag{5.15}$$

That is, each $\{|p\rangle\}$ picks up a complex exponential factor for its time evolution. It is immediately clear that $\langle P\rangle$ and $\langle(\Delta P)^2\rangle$ are independent of time. Calculationally, this occurs because P, and $(\Delta P)^2$ simplify to multiplication by numbers when acting on $|p\rangle$ states and the time-evolution complex-exponential factor cancels out because the two expectation values involve $\langle\psi|$ and $|\psi\rangle$. Physically, this occurs because the P operator commutes with H; later, we shall derive a general result about conservation of expectation values of operators that commute with the Hamiltonian. Either way one looks at it, one has

$$\langle P\rangle_t = \langle P\rangle_{t=0} = p_0 \qquad \langle(\Delta P)^2\rangle_t = \langle(\Delta P)^2\rangle_{t=0} = \sigma_p^2 \tag{5.16}$$

Let's also calculate the $\{|x\rangle\}$ representation of $|\psi(t)\rangle$. Here, we can just use our propagator formula, Equation 5.6, which tells us

$$\begin{aligned}
\psi_x(x,t) = \langle x\,|\psi(t)\rangle &= \int_{-\infty}^{\infty} dx'\,[U(t)]_{xx'}\,\langle x'\,|\psi(0)\rangle \\
&= \int_{-\infty}^{\infty} dx' \sqrt{\frac{m}{2\,\pi\,\hbar\,i\,t}}\; e^{\frac{i}{\hbar}\,\frac{m\,(x-x')^2}{2t}} \left(\frac{1}{2\,\pi\,\sigma_x^2}\right)^{1/4} e^{\frac{i}{\hbar}\,p_0\,x'}\, e^{-\frac{(x')^2}{4\,\sigma_x^2}} \\
&= \left[\sqrt{2\,\pi\,\sigma_x^2}\left(1+\frac{i\,\hbar\,t}{2\,m\,\sigma_x^2}\right)\right]^{-1/2} \exp\left[-\frac{\left(x-\frac{p_0}{m}\,t\right)^2}{4\,\sigma_x^2\left(1+\frac{i\,\hbar\,t}{2\,m\,\sigma_x^2}\right)}\right] \\
&\qquad \exp\left(\frac{i}{\hbar}\,p_0\,x\right) \exp\left(-\frac{i}{\hbar}\,\frac{p_0^2}{2\,m}\,t\right) \qquad (5.17)
\end{aligned}$$

where we do the integral in the usual fashion, by completing the square and using the Gaussian definite integral.

The Free Particle (cont.)

The probability density in the $\{|x\rangle\}$ basis is

$$P(x) = |\langle x|\psi(t)\rangle|^2$$
$$= \left[2\pi\sigma_x^2\left(1+\left(\frac{\hbar t}{2m\sigma_x^2}\right)^2\right)\right]^{-1/2} \exp\left[-\frac{\left(x-\frac{p_0}{m}t\right)^2}{2\sigma_x^2\left(1+\left(\frac{\hbar t}{2m\sigma_x^2}\right)^2\right)}\right] \quad (5.18)$$

Because the probability density is symmetric about $x = \frac{p_0}{m}t$, it is easy to see that

$$\langle X\rangle_t = \frac{p_0}{m}t = \langle X\rangle_{t=0} + \frac{p_0}{m}t \quad (5.19)$$

i.e., the particle's effective position moves with speed p_0/m, which is what one expects for a free particle with initial momentum p_0 and mass m.

The Free Particle (cont.)

The variance of the position is given by the denominator of the argument of the Gaussian exponential (one could verify this by calculation of the necessary integral),

$$\langle (\Delta X)^2 \rangle_t = \sigma_x^2 \left[1 + \left(\frac{\hbar t}{2 m \sigma_x^2} \right)^2 \right] = \langle (\Delta X)^2 \rangle_{t=0} \left[1 + \left(\frac{\hbar t}{2 m \sigma_x^2} \right)^2 \right] \tag{5.20}$$

The position uncertainty grows with time because of the initial momentum uncertainty of the particle – one can think of the $\{|p\rangle\}$ modes with $p > p_0$ as propagating faster than p_0/m and those with $p < p_0$ propagating more slowly, so the initial wavefunction spreads out over time. In the limit of large time ($t \gg 2 m \sigma_x^2/\hbar$), the uncertainty $\sqrt{\langle (\Delta X)^2 \rangle_t}$ grows linearly with time. The "large time" condition can be rewritten in a more intuitive form:

$$t \gg t_0 = 2 m \frac{\sigma_x^2}{\hbar} = m \frac{\sigma_x}{\sigma_p} = \frac{\sigma_x}{\sigma_v} \tag{5.21}$$

where $\sigma_v = \sigma_p/m$ is the velocity uncertainty derived from the momentum uncertainty. So, t_0 is just the time needed for the state with typical velocity to move the width of the initial state. We should have expected this kind of condition because σ_x and $\hbar$ are the only physical quantities in the problem. Such simple formulae can frequently be used in quantum mechanics to get quick estimates of such physical phenomena; we shall make such use in the particle in a box problem.

Position-Momentum Uncertainty Relation

Before leaving the free particle, we note an interesting relationship that appeared along the way. Recall that, because the position and momentum operators do not commute, $[X, P] = i\hbar$, no state is an eigenstate of both. If there is no uncertainty in one quantity because the system is in an eigenstate of it, then the uncertainty in the other quantity is in fact infinite. For example, a perfect position eigenstate has a delta-function position-space representation, but it then, by the alternative representation of the delta function, Equation 3.146, we see that it is a linear combination of all position eigenstates with equal weight. The momentum uncertainty will be infinite. Conversely, if a state is a position eigenstate, then its position-space representation has equal modulus everywhere and thus the position uncertainty will be infinite.

The Free Particle (cont.)

When we considered the Gaussian wave packet, which is neither an eigenstate of X nor of P, we found that the $t = 0$ position and momentum uncertainties were

$$\langle (\Delta X)^2 \rangle_{t=0} = \sigma_x^2 \qquad \langle (\Delta P)^2 \rangle_{t=0} = \sigma_p^2 = \frac{\hbar^2}{4\,\sigma_x^2}$$

Hence, at $t = 0$, we have the uncertainty relation

$$\sqrt{\langle (\Delta X)^2 \rangle_{t=0}} \sqrt{\langle (\Delta P)^2 \rangle_{t=0}} = \frac{\hbar}{2}$$

We saw that, for $t > 0$, the position uncertainty grows while the momentum uncertainty is unchanged, so in general we have

$$\sqrt{\langle (\Delta X)^2 \rangle} \sqrt{\langle (\Delta P)^2 \rangle} \geq \frac{\hbar}{2} \tag{5.22}$$

We will later make a general proof of this uncertainty relationship between noncommuting observables.

The Hamiltonian

A "box" consists of a region of vanishing potential energy surrounded by a region of infinite potential energy:

$$V(x) = \lim_{V_0 \to \infty} \begin{cases} 0 & |x| \leq \frac{L}{2} \\ V_0 & |x| > \frac{L}{2} \end{cases} \tag{5.23}$$

It is necessarily to include the limiting procedure so that we can make mathematical sense of the infinite value of the potential when we write the Hamiltonian. Classically, such a potential completely confines a particle to the region $|x| \leq L/2$. We shall find a similar result in quantum mechanics, though we need a bit more care in proving it.

The classical Hamiltonian is

$$\mathcal{H}(x, p) = \frac{p^2}{2\,m} + V(x) \tag{5.24}$$

The Particle in a Box (cont.)

Postulate 2 tells us that the quantum Hamiltonian operator is

$$H(X,P) = \frac{P^2}{2\,m} + V(X) \tag{5.25}$$

Next, we want to obtain an eigenvalue-eigenvector equation for H. For the free particle, when $V(X)$ was not present, it was obvious we should work in the $\{|p\rangle\}$ basis because H was diagonal there, and then it was obvious how P acted in that basis and we could write down the eigenvalues and eigenvectors of H trivially. We cannot do that here because $V(X)$ and hence H is not diagonal in the $\{|p\rangle\}$ basis. Moreover, regardless of basis, we are faced with the problem of how to interpret $V(X)$. Our usual power-series interpretation fails because the expansion is simply not defined for such a function – its value and derivatives all become infinite for $|x| \geq L/2$.

Shankar glosses over this issue and jumps to the final differential equation; thereby ignoring the confusing part of the problem! We belabor it to make sure it is clear how to get to the differential equation from H and the postulates. The only sensible way we have to deal with the above is to write down matrix elements of H in the $\{|x\rangle\}$ basis because our Postulate 2 tells us explicitly what the matrix elements of X are in this basis. Doing that, we have

$$\langle x\,|H(X,P)|x'\rangle = \langle x\,|\frac{P^2}{2\,m}\,|x'\rangle + \langle x\,|V(X)\,|x'\rangle \tag{5.26}$$

The Particle in a Box (cont.)

Let's look at each term separately. For the first term, since it is quadratic in P, let's insert completeness to get the P's separated:

$$\begin{aligned}
\langle x \,|\, \frac{P^2}{2\,m} \,|\, x' \rangle &= \frac{1}{2\,m} \int_{-\infty}^{\infty} dx'' \langle x \,|\, P \,|\, x'' \rangle \langle x'' \,|\, P \,|\, x' \rangle \\
&= -\frac{\hbar^2}{2\,m} \int_{-\infty}^{\infty} dx'' \left[\frac{d}{dx} \delta(x - x'') \right] \left[\frac{d}{dx''} \delta(x'' - x') \right] \\
&= -\frac{\hbar^2}{2\,m} \frac{d}{dx} \int_{-\infty}^{\infty} dx'' \delta(x - x'') \left[\frac{d}{dx''} \delta(x'' - x') \right] \\
&= -\frac{\hbar^2}{2\,m} \frac{d}{dx} \left[\frac{d}{dx} \delta(x - x') \right] \\
&= -\frac{\hbar^2}{2\,m} \delta(x - x') \frac{d^2}{d(x')^2} \qquad (5.27)
\end{aligned}$$

where in last step we used Equation 3.127,

$$\frac{d^n}{dx^n} \delta(x - x') = \delta(x - x') \frac{d^n}{dx'^n}$$

The Particle in a Box (cont.)

For the second term, we can approach it using a limiting procedure. Suppose $V(X)$ were not so pathological; suppose it has a convergent power series expansion $V(X) = \sum_{k=0}^{\infty} V_k X^k$. Then, we would have

$$\begin{aligned}\langle x\,|V(X)\,|x'\rangle &= \sum_{k=0}^{\infty} V_k \langle x\,|\,X^k|x'\rangle = \sum_{k=0}^{\infty} V_k \left(x'\right)^k \langle x\,|x'\rangle \\ &= \sum_{k=0}^{\infty} V_k \left(x'\right)^k \delta(x-x') = \delta(x-x')\,V(x')\end{aligned}$$

where we have allowed X to act to the right on $|x'\rangle$. This is not a strict application of Postulate 2; if one wants to be really rigorous about it, one ought to insert completeness relations like we did for P^2. For example, for X^2 we would have

$$\begin{aligned}\langle x\,|X^2\,|x'\rangle &= \int_{-\infty}^{\infty} dx''\langle x\,|X|x''\rangle\langle x''\,|X\,|x'\rangle = \int_{-\infty}^{\infty} dx''\,x\,\delta(x-x'')\,x''\,\delta(x''-x') \\ &= x^2\delta(x-x') = (x')^2\delta(x-x')\end{aligned}$$

For X^k, we have to insert $k-1$ completeness relations and do $k-1$ integrals. The result will be of the same form.

The Particle in a Box (cont.)

The key point in the above is that we have figured out how to convert the operator function $V(X)$ into a simple numerical function $V(x)$ when $V(X)$ can be expanded as a power series. To apply this to our non-analytic $V(X)$, we could come up with an analytic approximation that converges to the non-analytic one as we take some limit. (One could use a sum of $\tan^{-1}$ or tanh functions, for example.) The point is that if we used the expansion and then took the limit, we would obtain a result identical to the above. So we write

$$\langle x \,|V(X)\,| x' \rangle = \delta(x - x')\, V(x') \tag{5.28}$$

With the above results, we have that the matrix elements of H are given by:

$$\langle x \,|H\,| x' \rangle = \delta(x - x') \left[-\frac{\hbar^2}{2\,m}\, \frac{d^2}{d(x')^2} + V(x') \right] \tag{5.29}$$

The Particle in a Box (cont.)

Thus, for an arbitrary state $|f\rangle$, we have that (using completeness as usual)

$$\begin{aligned}
\langle x|H|f\rangle &= \int_{-\infty}^{\infty} dx' \langle x|H|x'\rangle\langle x'|f\rangle \\
&= \int_{-\infty}^{\infty} dx' \delta(x-x') \left[-\frac{\hbar^2}{2m}\frac{d^2}{d(x')^2} + V(x')\right] f(x') \\
&= \int_{-\infty}^{\infty} dx' \delta(x-x') \left[-\frac{\hbar^2}{2m}\frac{d^2 f(x')}{d(x')^2} + V(x') f(x')\right] \\
&= -\frac{\hbar^2}{2m}\frac{d^2 f(x)}{dx^2} + V(x) f(x) \qquad (5.30)
\end{aligned}$$

Lecture 16:
The One-Dimensional Particle in a Box Continued

Date Given: 2008/11/05
Date Revised: 2008/11/05

The Eigenvalue-Eigenvector Equation

Finally, we get to our eigenvalue-eigenvector equation for H, $E\,|\psi_E\,\rangle = H\,|\psi_E\,\rangle$. This equation can be written in the following form by applying $\langle x\,|$:

$$\langle x\,|E\,|\psi_E\,\rangle = \langle x\,|H\,|\psi_E\,\rangle = -\frac{\hbar^2}{2\,m}\,\frac{d^2}{dx^2}\,\psi_{E,x}(x) + V(x)\,\psi_{E,x}(x) \tag{5.31}$$

where $|\psi_E\,\rangle$ is the eigenstate of H corresponding to eigenvalue E (we wrote this as $|E\,\rangle$ earlier but that would prove confusing here) and $\psi_{E,x}(x) = \langle x\,|\psi_E\,\rangle$ is the $\{|x\,\rangle\}$-basis representation of $|\psi_E\,\rangle$. We rewrite:

$$\frac{d^2}{dx^2}\psi_{E,x}(x) + \frac{2\,m}{\hbar^2}\,[E - V(x)]\,\psi_{E,x}(x) = 0 \tag{5.32}$$

This is a second-order linear differential equation with the constant parameter E undetermined at this point; it will parameterize the solutions. We have thus reduced our eigenvalue-eigenvector problem to a differential equation. Solving this equation will give us the $\{|x\,\rangle\}$-basis representation of the eigenstates $|\psi_E\,\rangle$ and the allowed eigenvalues E.

The Particle in a Box (cont.)

That was a lot of work. By contrast, in the free particle case, we had to do no work because we already had a basis that was the eigenbasis of the Hamiltonian. Here, we have set up an equation that will give us the $\{|x\rangle\}$-basis representation of the eigenstates of H.

Finding the Eigenvectors

Our differential equation is a bit challenging because the $V(x)$ term is piecewise constant. Since it is straightforward to solve the differential equation for constant V, we will solve it separately in the three different regions $x < L/2$, $|x| \leq L/2$, and $x > L/2$ (which we will label I, II, and III), and then find conditions to make the three solutions consistent at the boundaries. We of course also keep V_0 finite for now. So we are searching for the solution to the generic equation

$$\frac{d^2}{dx^2}\psi_{E,x}(x) + \frac{2\,m}{\hbar^2}\,\alpha\,\psi_{E,x}(x) = 0$$

where $\alpha = E$ for $|x| \leq L/2$ and $\alpha = E - V_0$ for $|x| > L/2$. Because the coefficient of the $\psi_{E,x}(x)$ term is constant, the generic solution is an exponential,

$$\psi_{E,x}(x) = A\,e^{-\kappa\,x} + B\,e^{\kappa\,x} \qquad \kappa = \sqrt{-\frac{2\,m}{\hbar^2}\,\alpha}$$

where α may be positive or negative, and thus κ may be real or imaginary, depending on the value of E and $|x|$.

The Particle in a Box (cont.)

Let us consider states with $E < V_0$. (We don't care about $E \geq V_0$ states because we will let $V_0 \to \infty$ in the end.) For region I, the A term blows up as $x \to -\infty$ and must be discarded to keep the state normalizable; for region III, the B term similarly must be discarded. We therefore have

$$\psi^{\mathrm{I}}_{E,x}(x) = B_{\mathrm{I}}\, e^{\kappa\, x} \qquad \psi^{\mathrm{III}}_{E,x}(x) = A_{\mathrm{III}}\, e^{-\kappa\, x}$$

Next, let's consider our solution in region II. We rewrite

$$\psi^{\mathrm{II}}_{E,x}(x) = A\, e^{i\, k\, x} + B\, e^{-i\, k\, x} \qquad k = \sqrt{\frac{2\, m}{\hbar^2}\, E} \tag{5.33}$$

because we will see later that we obtain the requirement $E \geq 0$ and thus $\alpha \geq 0$. $E < 0$ solutions are not prevented by writing in the above form (they will simply have k imaginary), we'll just find the above form more convenient.

The Particle in a Box (cont.)

What are the requirements that will join the solutions in regions I, II, and III? Let us determine whether $\psi_{E,x}(x)$ must be continuous at the boundary by writing the change in $\psi_{E,x}(x)$ at the boundary in terms of its derivative:

$$\begin{aligned}
\psi_{E,x}(x)\Big|_{\frac{L}{2}-\epsilon}^{\frac{L}{2}+\epsilon} &= \int_{\frac{L}{2}-\epsilon}^{\frac{L}{2}+\epsilon} dx\, \frac{d}{dx}\, \psi_{E,x}(x) \\
&= \int_{\frac{L}{2}-\epsilon}^{\frac{L}{2}} dx\, \frac{d}{dx}\, \psi_{E,x}^{\mathrm{II}}(x) + \int_{\frac{L}{2}}^{\frac{L}{2}+\epsilon} dx\, \frac{d}{dx}\, \psi_{E,x}^{\mathrm{III}}(x) \\
&= \epsilon\, \frac{d}{dx}\, \psi_{E,x}^{\mathrm{II}}\left(\frac{L}{2}-\epsilon\right) + \epsilon\, \frac{d}{dx}\, \psi_{E,x}^{\mathrm{III}}\left(\frac{L}{2}+\epsilon\right) \\
&= i\,k\,\epsilon\left(A_{\mathrm{II}}\, e^{i\,k\,(\frac{L}{2}-\epsilon)} - B_{\mathrm{II}}\, e^{-i\,k\,(\frac{L}{2}-\epsilon)}\right) - \kappa\,\epsilon\, A_{\mathrm{III}}\, e^{-\kappa\,(\frac{L}{2}+\epsilon)}
\end{aligned}$$

where in, evaluating the integrals, we make use of the fact that the width of the interval will be taken to zero. This lets us take the value of the integrand at one endpoint and multiply by the width of the interval.

The Particle in a Box (cont.)

Now we take the limit $V_0 \to \infty$, which implies $\kappa \to \infty$. This has no effect on the first term. For the second term, we know $e^{-\kappa \frac{L}{2}} \to 0$ faster than $\kappa \to \infty$ because exponentials decay faster than any polynomial, so $\kappa\, e^{-\kappa \frac{L}{2}} \to 0$ as $V_0 \to \infty$.

Next we take $\epsilon \to 0$, which makes the first term vanish because all the quantities in it are finite.

So we find the $\psi_{E,x}$ must be continuous across the boundary.

Now, because $\kappa \to \infty$ as $V_0 \to \infty$, we also see that $\psi^{\mathrm{III}}_{E,x}(x) = 0$ identically in this limit. To make $\psi_{E,x}(x)$ continuous, we then have the requirement $\psi_{E,x}\left(\frac{L}{2}\right) = 0$.

The same condition holds at $x = -\frac{L}{2}$ by the same argument.

Do we need to find a condition on $\frac{d}{dx}\psi_{E,x}(x)$ at the boundaries? No. We have a second order differential equation and now we have two conditions on the solution, that it must vanish at both boundaries. That is enough initial/boundary data to fully determine the solutions. And, in fact, we will see that there is no condition on the derivative at the boundaries.

The Particle in a Box (cont.)

So, to make $\psi_{E,x}(x)$ vanish at $x = \pm L/2$, we have the following pair of equations that must be satisfied:

$$\begin{aligned} A\, e^{-i\,k\,\frac{L}{2}} + B\, e^{i\,k\,\frac{L}{2}} &= 0 \\ A\, e^{i\,k\,\frac{L}{2}} + B\, e^{-i\,k\,\frac{L}{2}} &= 0 \end{aligned} \tag{5.34}$$

where we have dropped the $_{\rm II}$ subscripts on A and B because they are no longer necessary. This is a pair of linear equations that can be written as a matrix equation

$$\begin{bmatrix} e^{-i\,k\,\frac{L}{2}} & e^{i\,k\,\frac{L}{2}} \\ e^{i\,k\,\frac{L}{2}} & e^{-i\,k\,\frac{L}{2}} \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The Particle in a Box (cont.)

We require the determinant vanish in order for there to be a nontrivial pair of coefficients A, B:

$$\begin{aligned} e^{-i\,k\,L} - e^{i\,k\,L} &= 0 \\ 2\,i\,\sin(k\,L) &= 0 \end{aligned} \tag{5.35}$$

which implies

$$k = \frac{n\,\pi}{L} \qquad n = 0, \pm 1, \pm 2, \ldots$$

Note that k may have no imaginary part if Equation 5.35 is to be satisfied. To see this, let's break k into real and imaginary parts, $k = k_r + i\,k_i$, and rewrite the above equation by moving the second term to the right side:

$$e^{-i\,k_r L} e^{k_i L} = e^{i\,k_r L} e^{-k_i L}$$

Now, take the modulus. The factors with k_r give 1, but the factors with k_i are always different than 1 and mismatched unless $k_i = 0$.

The Particle in a Box (cont.)

Now we must find the pairs A, B that satisfy the equation. Because the determinant of the above matrix vanishes, only one of the two linear equations is independent, so we write the first one using our newfound constraint on k:

$$A\,e^{-i\,n\,\frac{\pi}{2}} + B\,e^{i\,n\,\frac{\pi}{2}} = 0 \quad \Longrightarrow \quad A = -e^{i\,n\,\pi}\,B = (-1)^{n+1}B \tag{5.36}$$

The two complex exponentials are added with a $+$ or $-$ sign depending on whether n is odd or even.

The final result, then, is that we get a series of sines and cosines as the solutions:

$$\psi_{E,x}(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\,\pi\,x}{L}\right) \qquad n > 0, \text{even} \tag{5.37}$$

$$\psi_{E,x}(x) = \sqrt{\frac{2}{L}}\cos\left(\frac{n\,\pi\,x}{L}\right) \qquad n > 0, \text{odd}$$

We have dropped the negative values of n because they give the same functions up to possibly a sign flip, and we drop $n = 0$ because it gives nonsense for measurement of particle position ($P(x) = 0$ for all x).

The Particle in a Box (cont.)

Note that our solutions automatically satisfy the requirement we have for the K (and thus P) operator to be Hermitian, that the values at the two endpoints are equal, because our solutions vanish at the endpoints. The Hermiticity requirement can also be seen to be the more fundamental reason that k may have no imaginary part: if k had an imaginary part, the solution would include a real exponential, which is never the same at $-L/2$ and $L/2$.

Since the interval over which the wavefunction is nonzero will always be finite, we have been able to normalize the states to unity rather than to a delta function; the norm of a state is explicitly given by

$$|\psi_E|^2 = \langle \psi_E | \psi_E \rangle = \frac{1}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} dx \, \langle \psi_E | x \rangle \langle x | \psi_E \rangle = \frac{1}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} dx \, |\psi_{E,x}(x)|^2$$

Since the integral of $\sin^2$ or $\cos^2$ over any multiply of a half-wavelength is $1/2$, we thus see that our solutions are properly normalized. Orthogonality is also ensured by the properties of sines and cosines on this interval: one can check that $\langle \psi_E | \psi_{E'} \rangle = 0$ for $E \neq E'$, and hence we have $\langle \psi_E | \psi_{E'} \rangle = \delta_{EE'}$.

The Particle in a Box (cont.)

The eigenvalues are now found trivially by evaluating our differential equation, Equation 5.32, in region II (where $V(x)$ is finite and in fact vanishes):

$$E_n \psi_{E,x}(x) = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi_{E,x}(x) = \frac{\hbar^2 k_n^2}{2m}\psi_{E,x}(x)$$
$$\Longrightarrow \quad E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{\hbar^2 \pi^2 n^2}{2mL^2} \tag{5.38}$$

We see that the restriction to real k ensures that the energies can never be negative or imaginary. We thus see the requirement of Hermiticity of K return: Hermiticity of K ensures H is Hermitian, which ensures that the E_n will be real. The absence of $E < 0$ solutions does not come from Hermiticity but from physics: $E < 0$ states would be simply be physically unreasonable in this potential.

Quantization

We here see our first example of quantization of physical variables – here, the energy.

The quantization was due to the boundary conditions on the state's $\{|x\rangle\}$-space representation. Equation 5.33 for the solution in region II has only two degrees of freedom: the relative phases of the complex numbers A and B and the value of k. (The absolute magnitude of A and B would be set by normalization, and the absolute phase of A is arbitrary.) We found two conditions on this solution, also, based on the behavior of the state in regions I and II and on the continuity of the derivative at the boundaries. Two conditions and two degrees of freedom results in both degrees of freedom being restricted in some way. That results in discretization of k and the condition on the relative phase of A and B, $A = (-1)^{n+1}B$.

More fundamentally, the vanishing of the state in regions I and III arose from the fact that the potential is infinite in regions I and III, preventing the state from extending out into those regions. It is the inability of the particle to escape to infinity that results in quantization, as we will discuss more generally later.

In the free-particle case, we have all the trappings of quantum mechanics without the quantization because of the lack of such boundary conditions.

The Particle in a Box (cont.)

In the end, then, the thing that makes quantum mechanics "quantum" is not the postulates, including the fact that a state is a vector in a Hilbert space. Rather, it is the combination of these facts with the physical restrictions imposed by the Hamiltonian operator that results in quantization of physical variables, in the restriction of the eigenvalues to a discrete (if infinite) set.

However, it is important to remember that the other non-classical aspects of quantum mechanics – the information content in the wavefunction, the probabilistic treatment of physical observables including the ideas of expectation values and uncertainties – are certainly important and are present even when there is no explicit quantization.

The Nonzero Ground-State Energy

Why does the ground state have nonzero energy? Classically, we expect that the lowest energy state of a particle corresponds to no kinetic or potential energy. That state could easily be realized here, with the particle at rest anywhere in the box. Quantum mechanically, though, the lowest energy state has $E = \frac{\hbar^2}{2\,m}\,\frac{\pi^2}{L^2}$. Why is this?

This result is easily seen to be due to the noncommutativity of position and momentum. We saw earlier in the case of the free particle that position and momentum satisfy an uncertainty relation, Equation 5.22

$$\sqrt{\langle(\Delta X)^2\rangle}\sqrt{\langle(\Delta P)^2\rangle} \geq \frac{\hbar}{2}$$

Let us assume that this relation holds in general (we shall prove it later). The upper limit on the position uncertainty is $\sqrt{\langle(\Delta X)^2\rangle} \leq L/2$ because the particle's position-space probability distribution is nonzero only for $|x| \leq L/2$. Thus, we have

$$\sqrt{\langle(\Delta P)^2\rangle} \geq \frac{\hbar}{2\sqrt{\langle(\Delta X)^2\rangle}} \geq \frac{\hbar}{L}$$

The Particle in a Box (cont.)

Now, the particle energy is related to its momentum by $H = \frac{P^2}{2\,m}$ (we neglect $V(x)$ because it it vanishes everywhere that the wavefunction is nonzero). The same relation therefore holds for expectation values, $\langle H \rangle = \frac{\langle P^2 \rangle}{2\,m}$. Since $\langle P \rangle = 0$ for the ground state (and any eigenstate), we can thus obtain from our uncertainty relation a lower limit on $\langle H \rangle$:

$$\langle H \rangle = \frac{\langle P^2 \rangle}{2\,m} \geq \frac{1}{2\,m} \frac{\hbar^2}{2\,\langle (\Delta X)^2 \rangle} \geq \frac{\hbar^2}{2\,m\,L^2}$$

The actual value of the ground-state energy is π^2 bigger; had we done this more carefully by precisely calculating $\langle (\Delta X)^2 \rangle$, we would have gotten a closer match.

Regardless, we see that the nonzero ground-state energy is just a result of the nonzero ground-state momentum that arises from the containment of the particle in the interval $\left[-\frac{L}{2}, \frac{L}{2}\right]$. We shall see a similar result in the simple harmonic oscillator, giving rise to the famous $\frac{1}{2}\,\hbar\omega$ zero-point energy. It must be emphasized that this nonzero ground-state energy is not simply an offset that can be subtracted; it is real energy (purely kinetic for particle in a box, kinetic and potential for the harmonic oscillator) that has physical consequences. It affects, for example, the heat capacity of solids – even at absolute zero, there is vibrational motion of the ions in a crystal lattice.

Particle in a Box Propagator

We can use our generic formula Equation 4.18 to write the propagator that evolves forward any initial state:

$$U(t) = \sum_n e^{-\frac{i}{\hbar} E_n t} |\psi_{E_n}\rangle\langle\psi_{E_n}| = \sum_n \exp\left[-\frac{i}{\hbar}\frac{\hbar^2 \pi^2 n^2}{2 m L^2} t\right] |\psi_n\rangle\langle\psi_n| \tag{5.39}$$

where we use the shorthand $|\psi_n\rangle = |\psi_{E_n}\rangle$. In the $\{|x\rangle\}$-basis representation, we have

$$\begin{aligned}[U(t)]_{xx'} &= \langle x|U(t)|x'\rangle = \sum_n \exp\left[-\frac{i}{\hbar}\frac{\hbar^2 \pi^2 n^2}{2 m L^2} t\right] \langle x|\psi_n\rangle\langle\psi_n|x'\rangle \\ &= \sum_n \exp\left[-\frac{i}{\hbar}\frac{\hbar^2 \pi^2 n^2}{2 m L^2} t\right] \psi_{n,x}(x)\,\psi^*_{n,x}(x') \end{aligned} \tag{5.40}$$

where we abbreviate $\psi_{n,x}(x) \equiv \psi_{E_n,x}(x)$ for the position-space wavefunction for the nth energy level. This propagator cannot be further simplified, but it certainly provides a clear calculational method for time-evolving any initial state.

Lecture 17:
Relation of Particle in a Box Eigenstates to Position and Momentum Eigenstates
General Considerations on Bound States and Quantization
Continuity Equation for Probability

Date Given: 2008/11/07
Date Revised: 2008/11/07

Relation to $\{|x\rangle\}$ and $\{|p\rangle\}$ Basis States

We make the obvious point that our energy eigenstates are not position eigenstates: position operator eigenstates are independent of the Hamiltonian, and our states are just not the same as the position operator eigenstates in the $\{|x\rangle\}$-basis representation, so they are not the same in any representation.

We can also easily see that, even though the eigenstates of the Hamiltonian have a definite energy, they do not have a definite momentum. The momentum eigenstates are not dependent on the form of the Hamiltonian; it is always true that

$$\psi_{p,x}(x) = \langle x\,|p\rangle = \frac{1}{\sqrt{2\,\pi\,\hbar}}\,e^{\frac{i}{\hbar}\,p\,x} \tag{5.41}$$

Note that we consider momentum eigenstates for the entire real line because the inner product space we began with was for functions on the entire real line: while the wavefunction vanishes outside the box, it is still a well-defined function there. The position basis matrix representation $\psi_{E_n,x}(x) = \langle x\,|\psi_{E_n}\rangle$ of our eigenstate $|\psi_{E_n}\rangle$ is different from the above position-basis matrix representations of the $|p\rangle$ basis elements in two ways:

The Particle in a Box (cont.)

- Since the $\psi_{E_n,x}(x)$ are sines and cosines, we would need to take linear combinations of states at $|p\rangle$ and $|-p\rangle$ to obtain them.
- More importantly, $\psi_{E_n,x}(x)$ only equal to $\psi_{p,x}(x)$ over the interval $\left[-\frac{L}{2}, \frac{L}{2}\right]$, even though both $\psi_{E_n,x}(x)$ and $\psi_{p,x}(x)$ are defined on the entire real line.

Hence, our energy eigenstates are simply not equal to momentum eigenstates.

These facts are not surprising and arise simply from the fact that the Hamiltonian makes dynamics happen. In order for $\{|x\rangle\}$ basis states to be eigenstates of the Hamiltonian, a particle, once placed at a point, must never move from that point or spread out in position. Not even the free particle Hamiltonian allows that! And, while the free particle Hamiltonian's eigenstates are also momentum eigenstates, that clearly holds because there is no potential to affect the momentum of the particle. Once one includes any kind of potential, there is a force that can change the particle momentum and thus eigenstates of the Hamiltonian simply cannot be eigenstates of momentum.

Whence Quantization?

We made the point above that the quantization of energies for the particle in a box arises because of boundary conditions imposed by the potential energy function, not by the postulates. This argument holds generally for the bound states of any potential. We present a less detailed version of the argument given in Shankar Section 5.2 on this point.

Bound states are states whose energy E is less than the asymptotic value of the potential at $\pm\infty$. Classically, this is the case where the particle simply does not have enough energy to escape to $\pm\infty$. Quantum mechanically, the wavefunction must fall off at $\pm\infty$.

To make the argument, we need to count up the number of free parameters and see how they are determined by the boundary conditions. For an arbitrary potential, one can think of breaking it up into small intervals of size ϵ. As $\epsilon \longrightarrow 0$, the potential can be treated as piecewise constant. We thus have our exponential solutions in any interval, with the argument being imaginary or real depending on whether E is greater than or less than the value of $V(x)$ in the interval. There are four coefficient degrees of freedom for each of these intervals (the real and imaginary parts of the A and B coefficients).

General Considerations on Bound States and Quantization (cont.)

We have matching of both $\psi_x(x)$ and $\frac{d}{dx}\psi_x(x)$ at all the boundaries (the derivative must now match also because the steps in the potential are finite, as opposed to the particle-in-a-box case). That imposes four conditions (two equations each with real and imaginary parts) at each edge of the interval.

Now, let us cascade the conditions from left to right. Suppose the four coefficient degrees of freedom in the $-\infty$ interval have been set. That gives the four conditions that the wavefunction in the first finite interval must meet. The four coefficent degrees of freedom in the first finite interval thus are set. This procedure cascades through the last finite interval, which sets the four conditions at the last boundary. This provides enough information to set the four coefficient degrees of freedom in the $+\infty$ infinite interval. So, once the four coefficient degrees of freedom in the $-\infty$ interval and the energy E are set, the rest of the wavefunction is determined.

Now, let's consider how these first four coefficient degrees of freedom are set depending on whether we have a free or bound state.

General Considerations on Bound States and Quantization (cont.)

For a free state, we are allowed to keep both the $e^{\pm i\,k\,x}$ solutions since there is no worry about either blowing up at $\pm\infty$. That means we have both the A and B coefficients, and so we really have four degrees of freedom. For free states, we will always have two energy-degenerate states with the same energy, a right-going state and a left-going state. One can see this by noting that the energy eigenvalue of the Schrödinger Equation does not care about the sign of the argument of the imaginary exponential because two derivatives are taken:

$$\frac{d^2}{dx^2}e^{\pm i\,k\,x} = -k^2 e^{\pm i\,k\,x}$$

So, for a free particle, there will always be two independent, energy-degenerate solutions in the $\pm\infty$ regions. Since they are independent, their overall phases are arbitrary and independent. We make an arbitrary choice for this phase (*e.g.*, take A and B to be real in the $-\infty$ region) and that sets the phase of the two independent solutions. The amplitude of A and B in the $-\infty$ region will then be determined by normalization, as the two states must individually be normalized (to delta functions, since they do not decay away at $\pm\infty$). Once we have set A and B in the $-\infty$ region, the matching conditions cascade through to fully determine the wavefunction on the rest of the real line out to $+\infty$. Thus, in the end, we are able to meet all the conditions on the wavefunction using only the coefficient degrees of freedom. No quantization is required.

General Considerations on Bound States and Quantization (cont.)

For a bound state, the wavefunction must become real exponential outside of some $|x|$ for a given value of E because E is less than V somewhere: recall that the solution in any region where the potential is piecewise constant is

$$e^{\pm \kappa x} \qquad \text{with} \qquad \kappa = \sqrt{\frac{2\,m}{\hbar^2}\,(V-E)}$$

So, when $V > E$, κ is real.

Once we are forced into the real exponential solution, we acquire an additional condition we did not have in the free-particle case: the wavefunction may not blow up at $\pm\infty$ in order for it to be normalizable. This eliminates two of the four coefficient degrees of freedom in each infinite interval. Essentially, we get four additional conditions that we didn't have in the free particle case, which sets $A = 0$ in the $-\infty$ region and $B = 0$ in the $+\infty$ region. In the $-\infty$ interval, this presents no problem: rather than having two independent solutions, we have a single solution, but we still pick an arbitrary phase and normalization for the wavefunction in this region as before. Again, the matching conditions cascade through. However, now in the $+\infty$ region, instead of having four coefficient degrees of freedom to use to meet the four matching conditions, we only have the A coefficient degree of freedom because $B = 0$ is necessary to prevent the wavefunction from blowing up at $+\infty$. The only way to meet the four matching conditions is to restrict the freedom in E. Hence, we obtain energy quantization.

General Considerations on Bound States and Quantization (cont.)

The final case to consider is when the energy is such that the particle is bound on one side and free on the other. In such cases, there will be no energy quantization, but there will also not be independent left- and right-going solutions. The easiest way to see this is to begin defining the wavefunction on the bound side, which we will take to be the $-\infty$ side without loss of generality. The argument goes through as above, resulting in four matching conditions at the left edge of the $+\infty$ interval. Here, we can satisfy the four matching conditions without quantizing E simply be allowing the four matching conditions to set A and B in this region. A and B get tied together, tying the nominally independent $e^{\pm i\,k\,x}$ solution together. So, being bound on one side removes energy degeneracy but does not result in quantization. Classically, the way to see this is that, if you start out with a left-going state heading into this potential, it will encounter the potential barrier at $-\infty$ and be turned around into a right-going state. Since the eigenstates of the Hamiltonian are time-independent, they must contain both these behaviors. Thus, the solution is a superposition of left-going and right-going in the unbound region.

(As an aside, this is also a way to think about the solutions in the bound regions for which $E > V$: since the particle is bound and thus bounces off the potential barriers on the two sides, it can be in neither a pure left-going or right-going state. Thus, our solution in bound regions is always the sum of the left- and right-going waves.)

How did this all of this function in our particle in a box case? One can think of it in terms of keeping V_0 finite, considering only the bound state solutions, and then letting V_0 go to infinity. With finite V_0, the above generic bound state explanation would apply, resulting in energy quantization for bound states. Then, letting $V_0 \to \infty$ would not change this, but would simply eliminate any free states. We did not do the problem in this fashion because, by letting $V_0 \to \infty$ a bit earlier, we could conclude that the wavefunction vanished in the $\pm\infty$ regions. Rather than setting four conditions at the left side of the box using the wavefunction and its derivative, we instead set two conditions at each side of the both on the wavefunction alone. This was still four total conditions on the region II solution, fully specifying it. It would have been a bit more painful to do it in the generic way because we would have to carry along the non-vanishing region I and III solutions a bit longer, resulting in more algebra.

Analogy to Electromagnetism

Postulate 3 of QM tells us to interpret $|\langle x\,|\psi(t)\,\rangle|$ as the probability $P(x,t)$ that the position of the particle is in the interval x to $x+dx$ at time t; $P(x,t)$ is a probability density for finding the particle. This is similar to the idea of a charge density in electromagnetism, $\rho(x)$. In the case of electromagnetism, we have the following important results:

- The charge in an infinitesimal interval dx or volume d^3x is

$$dQ(x,t) = \rho(x,t)\,dx \quad \text{or} \quad dQ(\vec{x},t) = \rho(\vec{x},t)\,d^3x$$

- The total charge in an interval $[a,b]$ or volume V is

$$Q(t) = \int_a^b dx\,\rho(x,t) \quad \text{or} \quad Q(t) = \int_V d^3x\,\rho(\vec{x},t)$$

- The total charge over all space is conserved.

The Continuity Equation for Probability (cont.)

- The electric current density is defined as

$$j(x,t) = \rho(x,t)\, v(x,t) \quad \text{or} \quad \vec{j}(\vec{x},t) = \rho(\vec{x},t)\, \vec{v}(\vec{x},t)$$

where v or $\vec{v}$ is the velocity of the charges currently at position x or $\vec{x}$ at time t.
s

- The charge density satisfies the continuity equation

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial j(x,t)}{\partial x} = 0 \quad \text{or} \quad \frac{\partial \rho(\vec{x},t)}{\partial t} + \vec{\nabla} \cdot \vec{j}(x,t) = 0$$

Note the use of partial derivatives now. If one thinks of the charge density as a density smoothed over a large number of point charges, those charges are moving their position is a function of t. By $\frac{\partial}{\partial t}$, we mean "do not move along with the charge that moves at velocity v (or $\vec{v}$), just sit at a point x and watch the charges flow by and measure the rate of change of the density at the point x" and by $\frac{\partial}{\partial x}$, we mean "look at the gradient in x at a fixed point x, again do not move in x with the charges."

The Continuity Equation for Probability (cont.)

- The integral version of the continuity equation is

$$\frac{\partial}{\partial t}\int_a^b dx\,\rho(x,t) + [j(b,t) - j(a,t)] = 0$$

$$\text{or} \quad \frac{\partial}{\partial t}\int_V d^3x\,\rho(\vec{x},t) + \int_{S_V} d^2x\,\hat{n}(\vec{x})\cdot\vec{j}(\vec{x},t) = 0$$

  In the one-dimensional case, we initially had an integral over the interval $[a, b]$ of the perfect differential $\frac{\partial j(x,t)}{\partial x}$, which we simply integrated to get the difference of the values of $j(x, t)$ at the boundaries. In the three-dimensional case, we used Gauss's law to convert the volume integral of the divergence of $\vec{j}(\vec{x}, t)$ to a surface integral; $\hat{n}$ is the outward surface normal at $\vec{x}$. Note that, in both cases, the boundary is fixed in time.

The continuity equation says that charge must flow from one point to another in a smooth fashion – no sudden appearance or disappearance of charge is possible. Given that overall charge is conserved, we then have that the electrical current density must either vanish at infinity or the total current flux must vanish.

The Continuity Equation for Probability (cont.)

We can prove total conservation and the continuity equation for quantum-mechanical probability density, which is

$$P(x,t) = |\langle x\,|\psi(t)\,\rangle|^2 = |\psi_x(x,t)|^2 \quad \text{or} \quad P(\vec{x},t) = |\langle \vec{x}\,|\psi(t)\,\rangle|^2 = |\psi_x(\vec{x},t)|^2 \quad (5.42)$$

We shall see that the approprate definition for the associated probability current is

$$j(x,t) = -\frac{i}{2}\frac{\hbar}{m}\left(\psi_x^*(x,t)\frac{\partial}{\partial x}\psi_x(x,t) - \psi_x(x,t)\frac{\partial}{\partial x}\psi_x^*(x,t)\right) \quad (5.43)$$

$$\text{or} \quad \vec{j}(\vec{x},t) = -\frac{i}{2}\frac{\hbar}{m}\left(\psi_x^*(\vec{x},t)\vec{\nabla}\psi_x(\vec{x},t) - \psi_x(\vec{x},t)\vec{\nabla}\psi_x^*(\vec{x},t)\right)$$

The Continuity Equation for Probability (cont.)

First, let us prove that total probability is conserved. This is just a matter of using the fact that the Schrödinger Equation implies unitary evolution of the state when H is time-independent:

$$\begin{aligned}\int_V d^3x\, P(\vec{x},t) &= \int_V d^3x \langle\psi(t)\,|x\,\rangle\langle x\,|\psi(t)\,\rangle = \langle\psi(t)\,|\psi(t)\,\rangle \\ &= \langle\psi(0)\,|U^\dagger(t)U(t)|\psi(0)\,\rangle = \langle\psi(0)\,|\psi(0)\,\rangle\end{aligned}$$

Hence, if we take a time derivative of the integrated probability, we get zero: the total probability is conserved:

$$\frac{d}{dt}\int_V d^3x\, P(\vec{x},t) = 0 \tag{5.44}$$

.

The Continuity Equation for Probability (cont.)

Now, let us prove the other important result from E&M, the continuity equation. We prove this by explicitly taking the time derivative of the probability density using the Schrödinger Equation. We will prove this for the three-dimensional case; the restriction of the proof to one dimension will be clear. Note also that we immediately restrict to the position-space representation $\psi_x(\vec{x}, t) = \langle \vec{x} \,|\psi(t)\rangle$ because we are really only interested in the probability density and current in this representation. It would be interesting to discuss the probability current in other representations (momentum, energy, etc.), but is not relevant here.

The Continuity Equation for Probability (cont.)

We have

$$\begin{aligned}
\frac{\partial P(\vec{x},t)}{\partial t} &= \psi_x^*(\vec{x},t)\frac{\partial \psi_x(\vec{x},t)}{\partial t} + \psi_x(\vec{x},t)\frac{\partial \psi_x^*(\vec{x},t)}{\partial t} \\
&= \psi_x^*(\vec{x},t)\left[-\frac{i}{\hbar}H(\vec{x},t)\psi_x(\vec{x},t)\right] + \psi_x(\vec{x},t)\left[\frac{i}{\hbar}H^\dagger(\vec{x},t)\psi_x^*(\vec{x},t)\right] \\
&= -\frac{i}{\hbar}\,\psi_x^*(\vec{x},t)\left[-\frac{\hbar^2}{2\,m}\,\vec{\nabla}^2\psi_x(\vec{x},t) + V(\vec{x},t)\psi_x(\vec{x},t)\right] \\
&\quad + \frac{i}{\hbar}\,\psi_x(\vec{x},t)\left[-\frac{\hbar^2}{2\,m}\,\vec{\nabla}^2\psi_x^*(\vec{x},t) + V(\vec{x},t)\psi_x^*(\vec{x},t)\right] \\
&= \frac{i\,\hbar}{2\,m}\left[\psi_x^*(\vec{x},t)\vec{\nabla}^2\psi_x(\vec{x},t) - \psi_x(\vec{x},t)\vec{\nabla}^2\psi_x^*(\vec{x},t)\right]
\end{aligned}$$

where we used the Schrödinger Equation and its conjugate to get from the first line to the second, wrote out the Hamiltonian in the third line (using the fact that the potential $V(\vec{x},t)$ must be real for the Hamiltonian to be Hermitian and generalizing d^2/dx^2 to $\vec{\nabla}^2$ for three dimensions), and then canceled the $V(\vec{x},t)$ term to get the last line.

The Continuity Equation for Probability (cont.)

Finally, pull one $\vec{\nabla}$ to the front of the expression and manipulate a bit:

$$\begin{aligned}\frac{\partial P(\vec{x},t)}{\partial t} &= \vec{\nabla}\cdot\left[\frac{i\,\hbar}{2\,m}\left[\psi_x^*(\vec{x},t)\vec{\nabla}\psi_x(\vec{x},t)-\psi_x(\vec{x},t)\vec{\nabla}\psi_x^*(\vec{x},t)\right]\right] \\ &\quad -\left[\frac{i\,\hbar}{2\,m}\left[\left(\vec{\nabla}\psi_x^*(\vec{x},t)\right)\cdot\left(\vec{\nabla}\psi_x(\vec{x},t)\right)-\left(\vec{\nabla}\psi_x(\vec{x},t)\right)\cdot\left(\vec{\nabla}\psi_x^*(\vec{x},t)\right)\right]\right] \\ &= -\vec{\nabla}\cdot\vec{j}(\vec{x},t)\end{aligned}$$

and we have our desired result for the continuity equation:

$$\frac{\partial P(\vec{x},t)}{\partial t}+\vec{\nabla}\cdot\vec{j}(\vec{x},t)=0 \tag{5.45}$$

Lecture 18:
Scattering from a Step Potential

Date Given: 2008/11/10
Date Revised: 2008/11/10

Scattering from a Step Potential: Basic Solution

In this section, we are going to consider the problem of scattering of a particle from a step potential of height V_0. The complication here is that, physically, we want to start with incoming particles with well-defined position and momentum, send them into the barrier to interact, and calculate the outgoing particle state. But such initial and final states are not eigenstates of the Hamiltonian. So we will first solve for the eigenstates and then construct a wave-packet initial state, similar to what we did for the free particle.

We do things in a somewhat different order than Shankar, but the basic ideas are the same. Shankar begins with wave packets, which can be confusing since he a) has to fudge some things (*e.g.*, the packet width must be small compared to its initial distance from the step but large enough that the packet does not spread during the time it takes to reach the step) and b) considering the wave packet first violates our normal procedure of writing the Hamiltonian and finding the eigenstates first.

Setting up the Hamiltonian

We consider a potential

$$V(x) = \begin{cases} 0 & x < 0 \\ V_0 & x \geq 0 \end{cases} \tag{5.46}$$

We assume $V_0 \geq 0$ without lack of generality: if $V_0 < 0$, then one can reverse the x-axis to make the step orientation the same and adjust the solutions accordingly. The Hamiltonian is of course

$$H = \frac{P^2}{2\,m} + V(X)$$

As we did for the particle in a box, we assume the existence of an eigenstate of H of energy E, denoted by $|\psi_E\rangle$. The eigenvalue equation for H is identical in form to the one we wrote for the particle in a box except for the specific form of $V(x)$:

$$-\frac{\hbar^2}{2\,m}\frac{d^2}{dx^2}\psi_{E,x}(x) + V(x)\,\psi_{E,x}(x) = E\,\psi_{E,x}(x)$$

Solving for the Eigenstates, $E > V_0$

The stepwise nature of the potential suggests that we should solve the equation separately in the two regions: region I, $x < 0$; and region II, $x \geq 0$. In each region, the potential is constant, leading us to write the usual complex exponential form for the solution:

$$\psi^{\mathrm{I}}_{E,x}(x) = A\, e^{i\, k_1\, x} + B\, e^{-i\, k_1\, x} \qquad k_1 = \sqrt{\frac{2\, m}{\hbar^2}\, E} \tag{5.47}$$

$$\psi^{\mathrm{II}}_{E,x}(x) = C\, e^{i\, k_2\, x} + D\, e^{-i\, k_2\, x} \qquad k_2 = \sqrt{\frac{2\, m}{\hbar^2}\, (E - V_0)} \tag{5.48}$$

We restrict $k_1 \geq 0$ and $k_2 \geq 0$ because the sign-flip freedom in k_1 and k_2 is already allowed by having the two conjugate complex exponentials in each solution.

Since H is Hermitian, we are assured E is real and thus k_1 and k_2 are pure real or pure imaginary. Recall, though, that for H to be Hermitian, we need P to be Hermitian. When we consider functions on the entire real line, P is only Hermitian if the solutions either vanish or are exponentials with purely imaginary arguments at $\pm\infty$. For $E > V_0$, these conditions are all met without restricting the freedom in any of the coefficients. We will see that there will be restrictions when $E \leq V_0$.

Scattering from a Step Potential: Basic Solution (cont.)

As usual, to connect the solutions in the two regions, we need matching conditions. We showed even for the infinite well that the wavefunction must be continuous at the interface. The same result will hold here; the same proof goes through, and is in fact less subtle because everything stays finite. We can show that, in the case of a finite step, the first derivative of the wavefunction must also be continuous:

$$\begin{aligned}
\frac{d}{dx}\,\psi_{E,x}(x)\Big|_{-\epsilon}^{\epsilon} &= \int_{-\epsilon}^{\epsilon} dx\,\frac{d^2}{dx^2}\,\psi_{E,x}(x) \\
&= \frac{2\,m}{\hbar^2}\int_{-\epsilon}^{0} dx\,(-E)\;\psi^{\mathrm{I}}_{E,x}(x) + \frac{2\,m}{\hbar^2}\int_{0}^{\epsilon} dx\,(V_0 - E)\;\psi^{\mathrm{II}}_{E,x}(x) \\
&\approx \frac{2\,m}{\hbar^2}\,\epsilon\left[(-E)\psi^{\mathrm{I}}_{E,x}(-\epsilon) + (V_0 - E)\psi^{\mathrm{II}}_{E,x}(\epsilon)\right]
\end{aligned}$$

where we have used the eigenvalue equation to replace $\frac{d^2}{dx^2}\,\psi_{E,x}(x)$ and then rewritten the integrals assuming ϵ is small enough that ψ does not change appreciably over it. Since we have already shown $\psi_{E,x}(x)$ must be continuous across the boundary, we are guaranteed the difference in the parentheses will remain finite as $\epsilon \longrightarrow 0$. Thus, as we let $\epsilon \longrightarrow 0$, the entire expression is guaranteed to vanish, implying continuity of $\frac{d}{dx}\,\psi_{E,x}(x)$.

Scattering from a Step Potential: Basic Solution (cont.)

So our conditions for matching the wavefunction and its derivative are:

$$A + B = C + D \qquad i\,k_1\,(A - B) = i\,k_2\,(C - D) \tag{5.49}$$

We have four conditions (real and imaginary) and eight total degrees of freedom (real and imaginary of A, B, C, and D), so we have too much freedom. Recalling our discussion of the free particle and bound states, the excess freedom is in the fact that there are right-going and left-going states with the same energy; there is degeneracy for free states. Initially, one sees this from the fact that k_1 and k_2 could take on positive or negative values, though we restricted that freedom above, saying that it was already provided in the form of the solution. We apply this freedom here by making the ansatz that, for a given E, there should be two solutions, one that has $D = 0$ and one that has $A = 0$, corresponding to a right-going solution and a left-going solution. The right-going solution allows for a left-going solution in region I to allow for reflection at the interface, but no left-going solution in region II; and vice versa for the left-going solution. For each solution, we are left with four conditions and six degrees of freedom. The extra two are, as always, normalization and phase. *There is no need for quantization of k_1 or k_2 because there are enough coefficient degrees of freedom.*

Scattering from a Step Potential: Basic Solution (cont.)

Let's do the right-going solution explicitly; the left-going solution will be analogous. To eliminate the normalization and phase degrees of freedom explicitly, we take ratios, letting $b_\rightarrow = B/A$ and $c_\rightarrow = C/A$. We use $\rightarrow$ subscripts to indicate "right-going". So we have

$$1 + b_\rightarrow = c_\rightarrow \qquad\qquad k_1 (1 - b_\rightarrow) = k_2 c_\rightarrow$$
$$k_1 (1 - b_\rightarrow) = k_2 (1 + b_\rightarrow)$$
$$b_\rightarrow = \frac{k_1 - k_2}{k_1 + k_2} \qquad\qquad c_\rightarrow = \frac{2\,k_1}{k_1 + k_2} \tag{5.50}$$

The analogy to a wave on a string reaching an interface where the speed of propagation changes (due to a change in the string density) is clear. The situation is also similar to electromagnetic wave propagation across a boundary between two media with different indices of refraction. For the right-going solution, the left-going wave vanishes if $k_1 = k_2$, which corresponds to no step.

Scattering from a Step Potential: Basic Solution (cont.)

Before worrying about the time evolution, which, as we know, puts a simple time-dependent complex exponential in front, let's discuss the interpretation of these two states and their components. These solutions do not decay at infinity, so discussing the total probability is nonsense. We can, however, consider the probability currents that we defined in the previous section. Let us calculate them for the right-going state (*c.c.* =complex conjugate):

$$
\begin{aligned}
j^{\mathrm{I}}_{\rightarrow}(x) &= -\frac{i}{2}\,\frac{\hbar}{m}\left[\left(A^* e^{-i\,k_1\,x} + B^* e^{i\,k_1\,x}\right)\frac{\partial}{\partial x}\left(A\,e^{i\,k_1\,x} + B\,e^{-i\,k_1\,x}\right) - c.c\right] \\
&= -\frac{i}{2}\,\frac{\hbar}{m}\Big[\left(A^* e^{-i\,k_1\,x} + B^* e^{i\,k_1\,x}\right) i\,k_1\left(A\,e^{i\,k_1\,x} - B\,e^{-i\,k_1\,x}\right) \\
&\qquad\qquad - \left(A\,e^{i\,k_1\,x} + B\,e^{-i\,k_1\,x}\right) i\,k_1\left(-A^* e^{-i\,k_1\,x} + B^* e^{i\,k_1\,x}\right)\Big] \\
&= \frac{1}{2}\,\frac{\hbar\,k_1}{m}\Big[|A|^2 - |B|^2 + A\,B^*\,e^{2\,i\,k_1\,x} + A^*\,B\,e^{-2\,i\,k_1\,x} \\
&\qquad\qquad + |A|^2 - |B|^2 - A\,B^*\,e^{2\,i\,k_1\,x} - A^*\,B\,e^{-2\,i\,k_1\,x}\Big] \\
&= \frac{\hbar\,k_1}{m}\left[|A|^2 - |B|^2\right] = |A|^2\,\frac{\hbar\,k_1}{m}\left[1 - |b_{\rightarrow}|^2\right]
\end{aligned}
$$

Scattering from a Step Potential: Basic Solution (cont.)

The same calculation for region II is much easier because there is only one component to the wavefunction:

$$j_{\rightarrow}^{\mathrm{II}}(x) = -\frac{i}{2}\frac{\hbar}{m}\left[C^* e^{-i\,k_2\,x}\frac{\partial}{\partial x} C\, e^{i\,k_2\,x} - c.c\right]$$
$$= \frac{\hbar\,k_2}{m}\,|C|^2 = |A|^2\,\frac{\hbar\,k_2}{m}\,|c_{\rightarrow}|^2$$

When thinking in terms of probability currents, then, it becomes clear that the current breaks into three terms (where we now drop x because the currents have shown themselves to be constant in each region):

$$j_{\rightarrow}^{in,\mathrm{I}} = |A|^2\,\frac{\hbar\,k_1}{m} \tag{5.51}$$

$$j_{\rightarrow}^{out,\mathrm{I}} = -|A|^2\,\frac{\hbar\,k_1}{m}\,|b_{\rightarrow}|^2 \equiv -R\,j_{\rightarrow,\mathrm{I}}^{in} \qquad R = |b_{\rightarrow}|^2 \tag{5.52}$$

$$j_{\rightarrow}^{out,\mathrm{II}} = |A|^2\,\frac{\hbar\,k_2}{m}\,|c_{\rightarrow}|^2 \equiv T\,j_{\rightarrow,\mathrm{I}}^{in} \qquad T = |c_{\rightarrow}|^2\,\frac{k_2}{k_1} \tag{5.53}$$

Scattering from a Step Potential: Basic Solution (cont.)

The definitions of R and T ensure

$$R + T = |b_{\rightarrow}|^2 + \frac{k_2}{k_1}|c_{\rightarrow}|^2 = \left(\frac{k_1 - k_2}{k_1 + k_2}\right)^2 + \frac{k_2}{k_1}\left(\frac{2\,k_1}{k_1 + k_2}\right)^2 = 1 \tag{5.54}$$

which then implies

$$j_{\rightarrow}^{in,\mathrm{I}} + j_{\rightarrow}^{out,\mathrm{I}} = j_{\rightarrow}^{out,\mathrm{II}} \tag{5.55}$$

That is, we have conservation of the probability current at the step, as we expect.

The probability currents thus guide us in the interpretation of the states. The A term yields a right-going current in region I. The C term yields a right-going current in region II, which is clearly the part of the probability current that gets transmitted to region II. The B term yields a left-going current in region I, which we interpret as a reflected current. The interpretation of the B term as reflection is supported by the fact that $R = 1 - T$, which is consistent with conservation of probability current at the step.

Solving for the Eigenstates, $0 < E < V_0$

How do we solve the problem when $0 < E < V_0$? (We will consider $E < 0$ later.) Here, the allowed form for our solution is now

$$\psi^{\mathrm{I}}_{E,x}(x) = A\, e^{i\, k_1\, x} + B\, e^{-i\, k_1\, x} \qquad k_1 = \sqrt{\frac{2\, m}{\hbar^2}\, E} \tag{5.56}$$

$$\psi^{\mathrm{II}}_{E,x}(x) = C\, e^{-\kappa_2\, x} + D\, e^{\kappa_2\, x} \qquad \kappa_2 = \sqrt{\frac{2\, m}{\hbar^2}\, (V_0 - E)} \tag{5.57}$$

The solution for region I is the same as Equation 5.47, while the solution for region II has changed from a complex exponential in Equation 5.48 to a real exponential. The exponential constant κ_2 is related to the original version k_2 by $\kappa_2 = i\, k_2$. Again, we restrict $k_1 \geq 0$ and $\kappa_2 \geq 0$ because the sign flip freedom is already in the solution.

Since the solution is real exponential in region II, there is no longer a concept of right-going or left-going waves in that region. We shall see that we obtain a decaying *evanescent wave* in region II and that the probability current there vanishes.

Scattering from a Step Potential: Basic Solution (cont.)

Before continuing, we must set $D = 0$ to ensure the state is normalizable and hence belongs to the Hilbert space; left-going incoming states are thereby eliminated for this range of E and polarity of potential. One can also view this condition as being enforced by Hermticity of P, which requires that the wavefunction vanish or behave like an imaginary exponential at $\pm\infty$.

Our matching conditions then become

$$A + B = C \qquad i\,k_1\,(A - B) = -\kappa_2 C \tag{5.58}$$

The degrees of freedom now correspond to the $E > V_0$ case *after* we have picked either a left-going or right-going solution: there are four constraints and six complex coefficient degrees of freedom. This loss of degeneracy is because there is no left-going solution for $E < V_0$ and is consistent with our generic discussion of states that are unbound on one side. As before, two degrees of freedom go to phase and normalization, so we have a fully constrained problem *without any quantization requirement on* k_1 *or* κ_2.

Scattering from a Step Potential: Basic Solution (cont.)

Again, we set $b_{\rightarrow} = B/A$ and $c_{\rightarrow} = C/A$ and rewrite the above equations:

$$1 + b_{\rightarrow} = c_{\rightarrow} \qquad i\,k_1\,(1 - b_{\rightarrow}) = -\kappa_2 c_{\rightarrow}$$
$$i\,k_1\,(1 - b_{\rightarrow}) = -\kappa_2\,(1 + b_{\rightarrow})$$

$$b_{\rightarrow} = \frac{i\,k_1 + \kappa_2}{i\,k_1 - \kappa_2} = -\frac{(i\,k_1 + \kappa_2)^2}{k_1^2 + \kappa_2^2} = \frac{k_1^2 - \kappa_2^2}{k_1^2 + \kappa_2^2} - \frac{2\,i\,k_1\kappa_2}{k_1^2 + \kappa_2^2} \tag{5.59}$$

$$c_{\rightarrow} = \frac{2\,k_1^2}{k_1^2 + \kappa_2^2} - \frac{2\,i\,k_1\kappa_2}{k_1^2 + \kappa_2^2} \tag{5.60}$$

The coefficients are more complicated than in the $E > V_0$ case, but the matching procedure was the same. Notice that here $b_{\rightarrow}$ and $c_{\rightarrow}$ can pick up a phase shift, whereas in the $E > V_0$ case, they were real.

Scattering from a Step Potential: Basic Solution (cont.)

Let's calculate the probability currents again. The region I probability current keeps the same form because, though the formula for $b_{\rightarrow} = B/A$ has changed, the form of the solution is unchanged. On the right side, though, our functional form has changed and so we have to recalculate. The calculation is trivial though, now that the form on the right side is a real exponential. For this case, $\psi^{\mathrm{II}}_{E,x}(x)$ and $\psi^{\mathrm{II}*}_{E,x}(x)$ differ only in their coefficents, not in the arguments of the exponentials, so

$$\psi^{\mathrm{II}*}_{E,x}(x)\,\frac{d}{dx}\,\psi^{\mathrm{II}}_{E,x}(x) = \psi^{\mathrm{II}}_{E,x}(x)\,\frac{d}{dx}\,\psi^{\mathrm{II}*}_{E,x}(x)$$

(the conjugation of C does not matter because one gets $|C|^2$ in each term). Hence, the current on the right side vanishes! Moreover, one can easily see that $|b_{\rightarrow}| = 1$ because its numerator and denominator are complex conjugates of one another. To summarize, we have

$$j^{in}_{\rightarrow,\mathrm{I}} = |A|^2\,\frac{\hbar k_1}{m} \qquad j^{out}_{\rightarrow,\mathrm{I}} = -|A|^2\,\frac{\hbar k_1}{m} = -j^{in}_{\rightarrow,\mathrm{I}} \qquad j^{out}_{\rightarrow,\mathrm{II}} = 0 \tag{5.61}$$

$$R = 1 \qquad T = 0 \tag{5.62}$$

Scattering from a Step Potential: Basic Solution (cont.)

As one would expect classically for $E < V_0$, the right-going probability current in region II vanishes — the particle cannot escape to infinity on the right side because it does not have enough energy — and the reflected probability current is equal to the incoming probability current. However, there is a finite relative probability of measuring the particle's position to be in the barrier (relative to being in any finite interval in region I)! This is a completely quantum phenomenon that results from the fact that the particle state is a vector in a Hilbert space, not a definite position and momentum. If there step barrier were finite in length and dropped back to zero energy, we would find the probability current would not be zero in that region, indicating some probability for the incoming particle to tunnel through the barrier to a free-particle state on the right side. This is a good toy model for α decay in a nucleus.

Lecture 19:
Scattering from a Step Potential Continued: Eigenstates for Special Cases, Propagator, and Scattering of a Wave Packet

Date Given: 2008/11/12
Date Revised: 2008/11/12

Scattering from a Step Potential: Basic Solution

Special Cases: $E = V_0$, $E = 0$, and $E < 0$

- $E < 0$: The solutions must now be real exponential in both regions, and both the A and D coefficient terms must be eliminated to make the solution normalizable. The matching conditions now become impossible to meet:

$$B = C \qquad \kappa_1 B = -\kappa_2 C \tag{5.63}$$

$$\Longrightarrow \kappa_1 B = -\kappa_2 B \tag{5.64}$$

Since both κ_1 and κ_2 are positive numbers, one is left with $B = C = 0$. The solution becomes nonexistent, since the Hamiltonian acting on a state whose $\{|x\rangle\}$-basis representation vanishes everywhere returns zero, not a negative E.

- $E = 0$: Here, one must reconsider the differential equation in region I; it is now

$$-\frac{\hbar^2}{2\,m}\,\frac{d^2}{dx^2}\psi^{\mathrm{I}}_{E=0,x}(x) = 0$$

This is directly integrable; the general solution is

$$\psi^{\mathrm{I}}_{E=0,x}(x) = A + B\,x \tag{5.65}$$

We must exclude the second term because it blows up at $-\infty$. So we just have a constant in region I. This is ok, as a constant wavefunction is delta-function normalizable (we included $p = 0$ in our discussion of the free particle). The region II solution is the same as the $0 < E < V_0$ case because that part of the problem does not care about the value of E as long as $E < V_0$. The matching conditions become

$$A = C \qquad\qquad 0 = -\kappa_2 C \tag{5.66}$$

Since $\kappa_2 \neq 0$, we thus have $A = C = 0$. This ends up being the same as the $E < 0$ case.

Scattering from a Step Potential: Basic Solution (cont.)

- $E = V_0$: Here, the region II differential equation simplifies as the region I equation did for $E = 0$, so we have

$$\psi^{\mathrm{II}}_{E=V_0,x}(x) = C \tag{5.67}$$

(Again, the linear term must be discarded because it blows up at $+\infty$.) The solution in region I is the same as for $0 < E < V_0$ or $E > V_0$. Our matching conditions become

$$A + B = C \qquad i\,k_1\,(A - B) = 0 \tag{5.68}$$

So we have $A = B$ and $C = 2\,A$. This is a rather bizarre state: it is constant in region II. Note that, because it is constant, the momentum (derivative) there vanishes. The probability currents are simple and take on the same values that they do for the $0 < E < V_0$ case:

$$j^{in,\mathrm{I}}_{\rightarrow} = |A|^2\,\frac{\hbar\,k_1}{m} \qquad j^{out,\mathrm{I}}_{\rightarrow} = -|A|^2\,\frac{\hbar\,k_1}{m} = -j^{in,\mathrm{I}}_{\rightarrow} \qquad j^{out,\mathrm{II}}_{\rightarrow} = 0 \tag{5.69}$$

$$R = 1 \qquad T = 0 \tag{5.70}$$

Scattering from a Step Potential: Basic Solution (cont.)

The current vanishes not because the probability density vanishes, but because the state has no momentum in region II. But there is non-zero probability for finding the particle there. Here, our attempt to understand time-dependent behavior — incoming particles, transmission, reflection — using time-independent energy eigenstates breaks down. The probability currents still tell us something reasonable, but the wavefunction is not easily interpretable for time-dependent behavior. The right way to handle this is to construct a wave packet, which we will do next.

Propagator

Now that we have found the eigenstates, we can calculate the unitary propagator based on Equation 4.18

$$U(t) = \int_0^\infty dk\, e^{-\frac{i}{\hbar} E_k t} \left[|\psi_{E_k}\rangle\langle\psi_{E_k}| + \theta(-k_V - k)\, |\psi_{E_{-k}}\rangle\langle\psi_{E_{-k}}| \right] \tag{5.71}$$

where $|\psi_{E_k}\rangle$ is the state whose position representation has wavenumber $k_1 = k$ in region I. By analogy to the free particle case, we integrate over k (with an $\hbar$, we could have used p) instead of E to count states correctly and have the correct differential in the integral. Also, we have joined the integrals over negative k and over positive k by integrating only over positive k but summing the positive and negative k terms. Negative k corresponds to states that are initially left-going ($A = 0$, $D \neq 0$); the θ function is necessary because there are no such states for $0 < E < V_0$ for $V_0 > 0$. We define $k_V = \sqrt{\frac{2m}{\hbar^2} V_0}$, which is the wavevector of the state with energy V_0. The region II position wavefunction is specified by k also because

$$k_2 = \sqrt{\frac{2m}{\hbar^2}(E_k - V_0)} = \sqrt{k^2 - \frac{2m}{\hbar^2} V_0}$$

which picks up the appropriate factor of i when $E_k < V_0$.

Scattering from a Step Potential: Propagator (cont.)

The $\{|x\rangle\}$-basis matrix elements of $U(t)$ are given by taking the product with $\langle x|$ on the left and $|x\rangle$ on the right, giving

$$\begin{aligned}[U(t)]_{xx'} &= \int_0^\infty dk\, e^{-\frac{i}{\hbar}E_k t}\left[\langle x|\psi_{E_k}\rangle\langle\psi_{E_k}|x'\rangle + \theta(-k_V-k)\langle x|\psi_{E_{-k}}\rangle\langle\psi_{E_{-k}}|x'\rangle\right] \\ &= \int_0^\infty dk\, e^{-\frac{i}{\hbar}E_k t}\left[\psi^*_{E_k,x}(x)\,\psi_{E_k,x}(x') + \theta(-k_V-k)\,\psi^*_{E_{-k},x}(x)\,\psi_{E_{-k},x}(x')\right]\end{aligned} \tag{5.72}$$

Can we do the integral in closed form as we did for the free particle? Even for $x \leq 0$ and $x' \leq 0$, where the wavefunctions are free-particle-like, the integral cannot be done in the same way because of the absence of left-going states for $0 < E_k < V_0$. For $x > 0$ or $x' > 0$, the wavefunctions are either decaying exponentials or have an argument k_2 that is not related to E_k in the usual simple way, so the integral certainly cannot be done in the same way for them. So we leave the time-evolution operator in this form.

We note that we have not found the appropriate normalization factors A and D for our various eigenstates. We shall address this problem later. However, this means that the explicit form of the propagator's matrix elements remain undetermined up to a (possibly k-dependent!) normalizing factor.

Wave Packet Example

We have so far discussed eigenstates of the H and have interpreted the probability currents in such a way as to explain *time-dependent* behavior. This is a bit iffy. Let's do a time-dependent problem in an approximate form to substantiate this interpretation.

We will consider propagation of a Gaussian wave packet. Recall that such wave packet is defined in space at $t = 0$ by

$$\psi_x(x,0) = \langle x\,|\psi(0)\,\rangle = \left(\frac{1}{2\,\pi\,\sigma_x^2}\right)^{1/4} e^{\frac{i}{\hbar}\,p_0\,(x+a)} e^{-\frac{(x+a)^2}{4\,\sigma_x^2}} \tag{5.73}$$

with expectation values at $t = 0$ of

$$\langle X\rangle = -a \qquad \langle (\Delta X)^2\rangle = \sigma_x^2 \qquad \langle P\rangle = p_0 = \hbar\,k_0 \qquad \langle (\Delta P)^2\rangle = \sigma_p^2 = \frac{\hbar^2}{4\,\sigma_x^2} \tag{5.74}$$

Recall that our σ_x is related to Shankar's Δ_x by $\Delta_x = \sigma_x\sqrt{2}$.

Scattering from a Step Potential: Wave-Packet Example (cont.)

The wave packet is normalized to give unity probability in the initial state. Our idea is to time-evolve it forward, scattering it off the step potential, and then out to a large distance from the step. The packet will break into two separate pieces, one that goes off to $+\infty$ and one that goes off to $-\infty$. We will calculate the probability of transmission and reflection by calculating the total probability in each component, separately; this will be valid because they will have negligible overlap if evolved forward far enough in time.

Scattering from a Step Potential: Wave-Packet Example (cont.)

We want to have a wave packet that is localized enough that at $t = 0$ it does not have significant probability near the step so that we can really speak of the wave packet being essentially free before and after it interacts with the step. This implies

$$\sigma_x \ll a \qquad \Longleftrightarrow \qquad \frac{\sigma_x}{a} \ll 1 \tag{5.75}$$

We want the momentum to be well-defined so that the energy and the time needed to reach the barrier are well-defined. This requires

$$\sigma_p \ll p_0 \qquad \Longleftrightarrow \qquad \frac{\sigma_p}{p_0} \ll 1 \tag{5.76}$$

Finally, we want the time needed for the wave packet to approach and interact with the step potential to be small compared to the time over which the wave packet's width begins to increase appreciably so that the wave packet can be considered localized both before and after the interaction. Recall that the spreading time, Equation 5.21, is $t_0 = \frac{\sigma_x}{\sigma_v} = \frac{\sigma_x}{\sigma_p/m}$. The approach time is $t_{in} = \frac{a}{p_0/m}$. So we require

$$\frac{a}{p_0/m} \ll \frac{\sigma_x}{\sigma_p/m} \qquad \Longleftrightarrow \qquad \frac{\sigma_p}{p_0} \ll \frac{\sigma_x}{a} \tag{5.77}$$

These three conditions can be satisfied simultaneously.

Scattering from a Step Potential: Wave-Packet Example (cont.)

Because the momentum is so well defined and the wave packet is much smaller than the distance to the step (so the potential energy may be safely taken to be zero everywhere the wavefunction has significant amplitude), the initial energy is well defined. We can see this as follows:

$$\langle E\rangle = \frac{1}{2\,m}\langle P^2\rangle = \frac{1}{2\,m}\left(\langle P\rangle^2 + \langle(\Delta P)^2\rangle\right) = \frac{1}{2\,m}\left(p_0^2+\sigma_p^2\right) \approx \frac{p_0^2}{2\,m} \tag{5.78}$$

$$\begin{aligned}
\sigma_E^2 = \langle(\Delta E)^2\rangle &= \left(\frac{1}{2\,m}\right)^2\left(\langle P^4\rangle - \langle P^2\rangle^2\right) \\
&= \left(\frac{1}{2\,m}\right)^2\left(\langle P\rangle^4 + 6\langle P\rangle^2\langle(\Delta P)^2\rangle + 3\langle(\Delta P)^2\rangle^2 - \left(\langle P\rangle^2 + \langle(\Delta P)^2\rangle\right)^2\right) \\
&= \left(\frac{1}{2\,m}\right)^2\left(4\langle P\rangle^2\langle(\Delta P)^2\rangle + 2\langle(\Delta P)^2\rangle^2\right) = \left(\frac{1}{2\,m}\right)^2\left(4\,p_0^2\,\sigma_p^2 + 2\,\sigma_p^4\right)
\end{aligned}$$

$$\frac{\sigma_E}{E} = \frac{\sigma_p\sqrt{4\,p_0^2+2\,\sigma_p^2}}{p_0^2+\sigma_p^2} \approx 2\,\frac{\sigma_p}{p_0} \ll 1 \tag{5.79}$$

where we evaluated $\langle P^4\rangle$ using the method for the Gaussian integral $I_{2n}(\alpha)$ described in Shankar Appendix A.2, and in the last step we used $\sigma_p \ll p_0$.

Scattering from a Step Potential: Wave-Packet Example (cont.)

So, we have an initial state that satisfies all the conditions we have placed on it. To time-evolve it, we use the propagator we calculated above (Equations 5.71 and 5.72). The first step in applying the propagator is, as always, to take the matrix elements $\langle \psi_{E_k} | \psi(0) \rangle$. Let's first consider the inner product with the initially right-going states, which we labeled as $k > 0$. These are the states with $A \neq 0$, $D = 0$. We have

$$\langle \psi_{E_k} | \psi(0) \rangle = \int_{-\infty}^{\infty} dx \, \langle \psi_{E_k} | x \rangle \langle x | \psi(0) \rangle = \int_{-\infty}^{\infty} dx \, \psi^*_{E_k,x}(x) \, \psi_x(x,0)$$

$$= A_{\rightarrow}(k) \int_{-\infty}^{\infty} dx \, \left[\left(e^{i k x} + b_{\rightarrow}(k) \, e^{-i k x} \right) \theta(-x) + c_{\rightarrow}(k) \, e^{i k_2(k) x} \, \theta(x) \right]^* \psi_x(x,0)$$

where k is the wavevector in region I; $k_2(k)$ is the wavevector in region II as a function of k (and of course V_0); $b_{\rightarrow}(k)$ and $c_{\rightarrow}(k)$ are the coefficients defined earlier as a function of k_1 and k_2; since $k_1 = k$ and k_2 is a function of k for initially right-going states; we just write these as functions of k. We will leave the normalization undetermined until the end, writing it as $A_{\rightarrow}(k)$ to indicate it is the normalization of the right-going state at k. (We have taken it to be real without loss of generality.) The θ functions let us write our region I and II solutions together in concise form.

Scattering from a Step Potential: Wave-Packet Example (cont.)

Now, we may make some simplifications. By construction, the initial wavefunction $\psi_x(x,0)$ is essentially zero in region II ($x > 0$) where $\theta(x) \neq 0$; so the integral with the $c_{\rightarrow}$ term will vanish. By the same argument, we may drop $\theta(-x)$ as superfluous. Next, let's rewrite the above as follows:

$$\langle \psi_{E_k} | \psi(0) \rangle = A_{\rightarrow}(k) \int_{-\infty}^{\infty} dx\, e^{-i\,k\,x}\, \psi_x(x,0) + A_{\rightarrow}(k)\, b^*_{\rightarrow}(k) \int_{-\infty}^{\infty} dx\, e^{i\,k\,x}\, \psi_x(x,0)$$

where we have extended the integrals to be over the full real line because the vanishing of $\psi_x(x,0)$ in region II ensures the integral obtains no contribution from $x > 0$. The two terms are calculating the Fourier transform at k of $\psi_x(x,0)$ for positive and negative k, respectively. (Do not confuse this with the $A = 0, D \neq 0$ initially left-going states; we will come to those shortly. This is just a calculational analogy.)

Scattering from a Step Potential: Wave-Packet Example (cont.)

We know from our discussion of Gaussian wave packets for the free particle that the Fourier transform of our initial state is

$$\begin{aligned}\langle k\,|\psi(0)\,\rangle &= \frac{1}{\sqrt{2\,\pi}}\int_{-\infty}^{\infty} dx\, e^{-i\,k\,x}\left(\frac{1}{2\,\pi\,\sigma_x^2}\right)^{1/4} e^{i\,k_0\,(x+a)} e^{-\frac{(x+a)^2}{4\,\sigma_x^2}} \\ &= \left(\frac{1}{2\,\pi\,\sigma_k^2}\right)^{1/4} e^{-\frac{(k-k_0)^2}{4\,\sigma_k^2}}\, e^{i\,k\,a} \qquad \text{with} \quad \sigma_k = \frac{\sigma_p}{\hbar} = \frac{1}{2\,\sigma_x}\end{aligned}$$

where we have made the simple change of variables to $k = p/\hbar$ to match up with the way we index the states here with k. Since $\sigma_k/k_0 = \sigma_p/p_0 \ll 1$, it holds that the Gaussian essentially vanishes for $k < 0$. Therefore,

$$\langle \psi_{E_k}\,|\psi(0)\,\rangle \approx \sqrt{2\,\pi}\,A_{\rightarrow}(k)\left(\frac{1}{2\,\pi\,\sigma_k^2}\right)^{1/4} e^{-\frac{(k-k_0)^2}{4\,\sigma_k^2}}\, e^{i\,k\,a} \tag{5.80}$$

where the $\approx$ symbol indicates that this equality holds only for our approximations $\sigma_x/a \ll 1$, $\sigma_p/p_0 \ll 1$ and $\sigma_p/p_0 \ll \sigma_x/a$.

Scattering from a Step Potential: Wave-Packet Example (cont.)

If $E > V_0$, we must also calculate the the inner product of the initial state with the initially left-going states, which we labeled as $k < 0$. These are the states with $A = 0$, $D \neq 0$. (We don't need to worry about these if $0 < E \leq V_0$ because, recall, they don't exist for this range of energies.) We expect these inner products should essentially vanish because the initial state is almost entirely right-going. We have (skipping a step since it is identical to what we did before for the inner product with $|\psi_{E_k}\rangle$)

$$\langle \psi_{E_{-k}} | \psi(0) \rangle = \int_{-\infty}^{\infty} dx\, \psi^*_{E_{-k},x}(x)\, \psi_x(x,0)$$

$$= D_{\leftarrow}(k) \int_{-\infty}^{\infty} dx \left[\left(e^{-i\, k_2(k)\, x} + c_{\leftarrow}(k)\, e^{i\, k_2(k)\, x} \right) \theta(x) + b_{\leftarrow}(k)\, e^{-i\, k\, x}\, \theta(-x) \right]^* \psi_x(x,0)$$

Note that the role of b and c are exchanged here because the reflected and transmitted waves correspond to the C and B coefficients for an initially left-going wave (we did not write out explicit formulae for $c_{\leftarrow}(k)$ and $b_{\leftarrow}(k)$ earlier). We can see that all these terms vanish as follows. The first two terms vanish because the eigenstate wavefunction is only nonzero for $x > 0$ (the $\theta(x)$ factor), where, by construction, our initial state essentially vanishes. The third term vanishes because, even though it is nonzero for $x < 0$, it is a negative k wave; we argued above that the initial state has essentially no contribution from $k < 0$, so we may take that term to give zero also. So the initial state has no components in the initially left-going states, as we expected.

Since Equation 5.80 is thus the only term left for the initial state, we know that its probability density in k must integrate to unity. Therefore, we may conclude $A_{\rightarrow}(k) \approx 1/\sqrt{2\,\pi}$. Note the approximation symbol; we will obtain an exact value later and show how it reduces to this approximate result.

We have our inner products, so now the state at some time t later (for any $E > 0$) is

$$\begin{aligned}\langle x\,|\psi(t)\,\rangle &= \int_0^\infty dk\, e^{-\frac{i}{\hbar}\,E_k\,t}\langle x\,|\psi_{E_k}\,\rangle\langle\psi_{E_k}\,|\psi(0)\,\rangle \\ &\approx \int_0^\infty dk\, e^{-\frac{i}{\hbar}\,E_k\,t}\,\frac{1}{\sqrt{2\,\pi}}\Big[\left(e^{i\,k\,x} + b_{\rightarrow}(k)\,e^{-i\,k\,x}\right)\theta(-x) + c_{\rightarrow}(k)\,e^{i\,k_2(k)\,x}\,\theta(x)\Big] \\ &\qquad \times \left(\frac{1}{2\,\pi\,\sigma_k^2}\right)^{1/4} e^{-\frac{(k-k_0)^2}{4\,\sigma_k^2}}\; e^{i\,k\,a}\end{aligned}$$

For $0 < E \leq V_0$, the $c_{\rightarrow}$ term goes from a complex exponential to a real exponential as necessary. Our arguments using the vanishing of the initial state for $x > 0$ and $k < 0$ continue to hold to give us the same Gaussian form for $\langle\psi_{E_k}\,|\psi(0)\,\rangle$ even when $E < V_0$.

Scattering from a Step Potential: Wave-Packet Example (cont.)

We note that $b_{\rightarrow}(k)$ and $c_{\rightarrow}(k)$ are slowly varying functions of k compared to the Gaussian in k that is sharply peaked at $k = k_0$, so we may approximate them as constant at $k = k_0$ and pull them out of the integral. We also separate the terms:

$$\begin{aligned}
\langle x\,|\psi(t)\,\rangle = {} & \frac{\theta(-x)}{\sqrt{2\,\pi}} \int_0^\infty dk\, e^{-\frac{i}{\hbar}\,E_k\,t} e^{i\,k\,x} \left(\frac{1}{2\,\pi\,\sigma_k^2}\right)^{1/4} e^{-\frac{(k-k_0)^2}{4\,\sigma_k^2}}\, e^{i\,k\,a} \\
& + \frac{\theta(-x)\,b_{\rightarrow}(k_0)}{\sqrt{2\,\pi}} \int_0^\infty dk\, e^{-\frac{i}{\hbar}\,E_k\,t} e^{-i\,k\,x} \left(\frac{1}{2\,\pi\,\sigma_k^2}\right)^{1/4} e^{-\frac{(k-k_0)^2}{4\,\sigma_k^2}}\, e^{i\,k\,a} \\
& + \frac{\theta(x)\,c_{\rightarrow}(k_0)}{\sqrt{2\,\pi}} \int_0^\infty dk\, e^{-\frac{i}{\hbar}\,E_k\,t} e^{i\,k_2(k)\,x} \left(\frac{1}{2\,\pi\,\sigma_k^2}\right)^{1/4} e^{-\frac{(k-k_0)^2}{4\,\sigma_k^2}}\, e^{i\,k\,a}
\end{aligned}$$

Let's check this form at $t = 0$. The first term is the positive k part of the Fourier reverse transform integral. Since the $\{|k\,\rangle\}$-basis representation of the original state is sharply peaked at $k_0 > 0$ with $\sigma_k/k_0 \ll 1$, this positive k part of the integral is essentially the full integral. The second term is $b_{\rightarrow}(k_0)$ times the negative k part of the integral. The negative k part of the integral contributes negligibly to the Fourier transform, so that term can be taken to vanish. The third term is trickier. But we don't need to evaluate it directly. Since the first term is essentially the initial state, all the probability is there and the third term must vanish by simple conservation of probability.

Scattering from a Step Potential: Wave-Packet Example (cont.)

Now, let's evaluate $\langle x\,|\psi(t)\,\rangle$. To determine the reflected and transmitted probabilities, we want to know what happens well after the interaction with the step potential, $t \sim t_{out} = 2\,t_{in} = 2\,\frac{a}{p_0/m}$.

The first term looks like the positive k part of the free-particle propagator acting on the initial state. Because the initial state is well-defined in momentum about $k_0 > 0$, this is a very good approximation to the entire propagator. That propagator would move the wave packet forward with speed p_0/m, which would leave it centered at $-a + \frac{p_0}{m}\,t$ at time t; for $t \sim t_{out}$, the wave packet would thus be well to positive x. We showed earlier that the wave packet does not spread very quickly; for $t \sim t_{out}$, the spreading remains negligible compared to its position. So, for such times, this wave packet would thus have almost vanishing value for $x < 0$. For later times, the spreading becomes linear in time; but, because the propagation is linear in time, the value of the wavefunction at $x < 0$ remains negligible. So, since this term is multiplied by $\theta(-x)$, the entire term vanishes for $t \sim t_{out}$ and later.

Scattering from a Step Potential: Wave-Packet Example (cont.)

To evaluate the second term, let's change the integration variable from k to $k' = -k$:

$$\text{2nd term} = \frac{\theta(-x)\, b_{\rightarrow}(k_0)}{\sqrt{2\,\pi}} \int_{-\infty}^{0} dk'\, e^{-\frac{i}{\hbar} E_k\, t} e^{i\, k'\, x} \left(\frac{1}{2\,\pi\,\sigma_k^2}\right)^{1/4} e^{-\frac{(k'+k_0)^2}{4\,\sigma_k^2}}\, e^{-i\, k'\, a}$$

We now have the negative k part of the integral for the free-particle propagation of a state centered tightly around momentum $-k_0$ and that started at position $x = +a$ at $t = 0$ (*i.e.*, in region II). By the same kind of argument as we used above for the first term, this is essentially the entire free-particle propagator for an initial state with position $a > 0$, mean momentum $p_0 = -\hbar\, k_0$, and momentum spread $\sigma_p = \hbar\,\sigma_k << p_0$. The $\theta(-x)$ function is now 1 for $t \sim t_{out}$ because the state will be at $x \sim -a$ at that time, so the θ function does nothing. So we have

$$\text{2nd term} = \left[\frac{1}{2\,\pi\,\sigma_x^2\left(1 + \frac{i\,\hbar\,t}{2\,m\,\sigma_x^2}\right)}\right]^{-1/4} \exp\left[-\frac{\left(x + \frac{\hbar\, k_0}{m}\, t\right)^2}{4\,\sigma_x^2\left(1 + \frac{i\,\hbar\,t}{2\,m\,\sigma_x^2}\right)}\right] e^{-i\, k_0\, x} e^{-\frac{i}{\hbar} E_{k_0}\, t}$$

This is just a left-going Gaussian wave packet with mean position $x(t) = a - \frac{p_0}{m}\, t$, mean (and constant) momentum $-p_0$, constant momentum width σ_p, and a position width that spreads with time in the same way as for the free particle. That is, this term looks like a free particle that propagated from $x > 0$ with no effect from the barrier.

The third term is much harder to evaluate. We make a generic argument to obtain the probabilities in the reflected and transmitted waves. The second term is a standard Gaussian multiplied by $b_{\rightarrow}(k_0)$. The Gaussian will just give unity when the probability is calculated, so the probability in the reflected part is

$$R = |b_{\rightarrow}(k_0)|^2 \tag{5.81}$$

By conservation of probability, we know that the probability in the transmitted wave is

$$T = 1 - R \tag{5.82}$$

We know that $|b_{\rightarrow}(k_0)| = 1$ for $0 < E < V_0$ by our direct calculation of $b_{\rightarrow}(k)$ earlier (Equations 5.50, 5.59, and 5.60), so we are even assured that $R = 1$ and $T = 0$ for $0 < E < V_0$: the probability at large time is completely reflected by the barrier.

Lecture 20:
Scattering from a Step Potential Continued:
Normalization of Eigenstates
Date Given: 2008/11/14
Date Revised: 2008/11/14

Normalizing the Step Potential Eigenstates

The attentive reader will notice that we never calculated an exact for form for $A_{\rightarrow}(k)$, the normalization of the right-going eigenstates for the step potential. Let's return to that calculation and check that the result reduces to $A_{\rightarrow}(k) = 1/\sqrt{2\,\pi}$, the approximate value we determined above using the wave-packet example.

Let's first do the calculation for $E \geq V_0$ in the obvious fashion:

$$\begin{aligned}
\langle \psi_{E_k} | \psi_{E_k} \rangle &= \int_{-\infty}^{\infty} dx \, |\psi_{E_k,x}(x)|^2 \\
&= |A_{\rightarrow}|^2 \left[\int_{\infty}^{0} dx \, \left[1 + |b_{\rightarrow}|^2 + b_{\rightarrow} \left(e^{-2\,i\,k\,x} + e^{2\,i\,k\,x} \right) \right] + \int_{0}^{\infty} dx \, |c_{\rightarrow}|^2 \right] \\
&= |A_{\rightarrow}|^2 \left[\left(1 + |b_{\rightarrow}|^2 + |c_{\rightarrow}|^2 \right) \frac{1}{2} \int_{-\infty}^{\infty} dx + b_{\rightarrow} \int_{-\infty}^{\infty} dx \, e^{2\,i\,k\,x} \right] \\
&= |A_{\rightarrow}|^2 \left[\left(1 + |b_{\rightarrow}|^2 + |c_{\rightarrow}|^2 \right) \pi \langle k | k \rangle + 2\,\pi \langle -k | k \rangle \right] \\
&= |A_{\rightarrow}|^2 \, \pi \, \left(1 + |b_{\rightarrow}|^2 + |c_{\rightarrow}|^2 \right) \langle k | k \rangle
\end{aligned}$$

In the first step, we make use of the fact that, for $E_k > V_0$, the wavefuncion in region II has unity modulus. In the second step, we use $\int_{-\infty}^{0} dx = \frac{1}{2} \int_{-\infty}^{\infty} dx$ and change variables on $\int_{-\infty}^{0} dx \, e^{-2\,i\,k\,x}$ so that it becomes $\int_{0}^{\infty} dx \, e^{2\,i\,k\,x}$. In the next step, we recognize that these integrals are what one gets when one calculates inner products of $\{|k\rangle\}$ basis elements, and then that $\langle -k | k \rangle = 0$ by orthogonality.

Scattering from a Step Potential: Normalization of Eigenstates (cont.)

Next, we need to decide what we want $\langle \psi_{E_k} | \psi_{E_k} \rangle$ to be so we can solve for $|A_{\rightarrow}|^2$. Since our eigenstates for $E_k \geq V_0$ extend to $\pm\infty$, they must be delta-function normalized, so presumably we want $\langle \psi_{E_k} | \psi_{E_k} \rangle = \delta(0) = \infty$. And it is true that $\langle k | k \rangle = \delta(0) = \infty$. However, having ∞ on both sides of an equation can result in missing factors because any number times ∞ is ∞. Let's be a bit more careful.

Since Equation 5.71 used dk as the differential for the integral over all possible states, we want in general to have $\langle \psi_{E_k} | \psi_{E_{k'}} \rangle = \delta(k - k')$. This is the same as the normalization of the $\{|k\rangle\}$ basis, $\langle k | k' \rangle = \delta(k - k')$. We are now assured the $\langle \psi_{E_k} | \psi_{E_k} \rangle = \langle k | k \rangle$ and there is no issue of a missing scale factor between the two. There would have been had we chosen, for example $\hbar\, dk$ as the measure for the differential. In any case, we are now free to cancel $\delta(0)$ from both sides, leaving:

$$1 = |A_{\rightarrow}|^2 \pi \left(1 + |b_{\rightarrow}|^2 + |c_{\rightarrow}|^2\right)$$

We have the freedom to choose $A_{\rightarrow}$ to be real, so we may solve the above:

$$\begin{aligned} A_{\rightarrow}(k) &= \frac{1}{\sqrt{2\pi}} \left(\frac{2}{1 + |b_{\rightarrow}(k)|^2 + |c_{\rightarrow}(k)|^2} \right)^{1/2} \\ &= \frac{1}{\sqrt{2\pi}} \left(1 + \frac{2k(k - k_2)}{(k + k_2)^2} \right)^{-1/2} \qquad E_k \geq V_0 \end{aligned} \tag{5.83}$$

Now, interestingly, when $E_k \geq V_0$, the only limit in which the above expression reduces to $A \approx 1/\sqrt{2\,\pi}$ is the limit $E_k \gg V_0$ because then $k_2 \approx k$ and the second term vanishes. So, surprisingly, we discover another condition that must be imposed in order for our wave packet solution to be valid: $E \gg V_0$ if $E > V_0$. The reason this is necessary but was not apparent initially is that it is a condition that derives from the transmitted wave packet in region II. Essentially, in order for our wave packet approximation to be valid, we should have placed the additional condition that $\sigma_k/k_2 \ll \sigma_x/a$: this ensures that the transmitted wave packet on the right side also meets our requirements about being well-defined in x, in p, and to not spread too quickly.

It is amusing to note that, when $E_k = V_0$ and $k_2 = 0$, the above reduces to

$$A_{\rightarrow}(k = k_V) = \frac{1}{\sqrt{2\,\pi}} \left(1 + \frac{2\,k\,(k-0)}{(k+0)^2}\right)^{-1/2} = \frac{1}{\sqrt{3}\,\sqrt{2\,\pi}}$$

Scattering from a Step Potential: Normalization of Eigenstates (cont.)

How does this normalization issue work out for $E_k < V_0$? In this case, the wavefunction in region II is a decaying real exponential and its contribution to the normalization integral is finite and can thus be ignored compared to the infinite contribution from region I. So, we have

$$A_{\rightarrow}(k) = \frac{1}{\sqrt{2\,\pi}} \left(\frac{2}{1 + |b_{\rightarrow}(k)|^2} \right)^{1/2} = \frac{1}{\sqrt{2\,\pi}} \left(\frac{2}{1 + R} \right)^{1/2}$$
$$= \frac{1}{\sqrt{2\,\pi}} \qquad 0 < E_k < V_0 \tag{5.84}$$

Our approximate solution for $A_{\rightarrow}$ is exact when $E < V_0$ and the wave packet is completely reflected. Note that the condition $\sigma_p/p_0 \ll 1$ implies $\sigma_E/E \ll 1$ and thus $\sigma_E/V_0 \ll 1$ in this case: the wave packet energy width must be small compared to the potential height. This ensures that none of the momentum modes in the wave packet can be transmitted.

Theorems on One-Dimensional States

There is no degeneracy in one-dimensional bound states.

The proof is given in Shankar; we repeat it here only for completeness.

Suppose there are two degenerate bound states with the same energy E, $|\psi_{E,1}\rangle$ and $|\psi_{E,2}\rangle$. By "bound", we mean that their $\{|x\rangle\}$-basis representations, $\psi_{E,1,x}(x) = \langle x|\psi_{E,1}\rangle$ and $\psi_{E,2,x}(x) = \langle x|\psi_{E,2}\rangle$, fall off at $\pm\infty$ faster than $1/\sqrt{|x|}$ so that they can be normalized to unity rather than a delta function. Then they both are eigenvectors of H with the same eigenvalue E: $H|\psi_{E,1}\rangle = E|\psi_{E,1}\rangle$ and $H|\psi_{E,2}\rangle = E|\psi_{E,2}\rangle$. The $\{|x\rangle\}$-basis representation version of these statements is

$$\langle x|H|\psi_{E,1}\rangle = \langle x|E|\psi_{E,1}\rangle \qquad \langle x|H|\psi_{E,2}\rangle = \langle x|E|\psi_{E,2}\rangle$$

We may write the above as (recall, use the $\{|x\rangle\}$-basis matrix elements of H and insert completeness in this basis):

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi_{E,1,x}(x) + V(x)\psi_{E,1,x}(x) = E\psi_{E,1,x}(x)$$
$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi_{E,2,x}(x) + V(x)\psi_{E,2,x}(x) = E\psi_{E,2,x}(x)$$

Theorems on One-Dimensional States (cont.)

Multiply the first equation by $\psi_{E,2,x}(x)$ and the second by $\psi_{E,1,x}(x)$ and difference to obtain

$$\psi_{E,1,x}(x)\,\frac{d^2}{dx^2}\,\psi_{E,2,x}(x) - \psi_{E,2,x}(x)\,\frac{d^2}{dx^2}\,\psi_{E,1,x}(x) = 0$$

$$\frac{d}{dx}\left(\psi_{E,1,x}(x)\,\frac{d}{dx}\,\psi_{E,2,x}(x) - \psi_{E,2,x}(x)\,\frac{d}{dx}\,\psi_{E,1,x}(x)\right) = 0$$

$$\psi_{E,1,x}(x)\,\frac{d}{dx}\,\psi_{E,2,x}(x) - \psi_{E,2,x}(x)\,\frac{d}{dx}\,\psi_{E,1,x}(x) = c$$

where in the last step we simply integrated the total differential. Now, assuming that $\psi_{E,1,x}(x)$ and $\psi_{E,2,x}(x)$ are bound states, they must vanish as $|x| \to \infty$. Thus, the left side of the equation vanishes as $|x| \to \infty$; since the right side is constant, $c = 0$ is needed.

Theorems on One-Dimensional States (cont.)

With $c = 0$, we have

$$\frac{1}{\psi_{E,1,x}(x)} \frac{d}{dx} \psi_{E,1,x}(x) = \frac{1}{\psi_{E,2,x}(x)} \frac{d}{dx} \psi_{E,2,x}(x)$$

$$\frac{d}{dx} \left[\ln \psi_{E,1,x}(x) \right] = \frac{d}{dx} \left[\ln \psi_{E,2,x}(x) \right]$$

$$\ln \psi_{E,1,x}(x) = \ln \psi_{E,2,x}(x) + d$$

$$\psi_{E,1,x}(x) = e^d \psi_{E,2,x}$$

d is a constant, so the two states are the same up to a constant – hence, they are the same state. In fact, if they had both been unity-normalized, then $d = 0$ is necessary and they are identically functions, too.

Theorems on One-Dimensional States (cont.)

The eigenfunctions of $H = \frac{P^2}{2\,m} + V(X)$ in one dimension can always be chosen to be real.

Again, the proof is given in Shankar; we repeat it here only for completeness. We also try to explain better under what conditions it fails.

By the same arguments as given for the previous theorem, consider the eigenvalue equation for H and the state ψ_E with eigenvalue E:

$$-\frac{\hbar^2}{2\,m}\,\frac{d^2}{dx^2}\,\psi_{E,x}(x) + V(x)\,\psi_{E,x}(x) = E\,\psi_{E,x}(x)$$

For Hamiltonians of the above form, we can complex conjugate the above equation:

$$-\frac{\hbar^2}{2\,m}\,\frac{d^2}{dx^2}\,\psi^*_{E,x}(x) + V(x)\,\psi^*_{E,x}(x) = E\,\psi^*_{E,x}(x)$$

We see that $\psi^*_{E,x}(x)$ is the $\{|x\rangle\}$-basis representation of a state that is also an eigenstate of *the same* H with the same eigenvalue E.

Theorems on One-Dimensional States (cont.)

The fact that H stayed the same under complex conjugation is a result of its particular form; if one had $H = \frac{P^2}{2\,m} - \beta\,\frac{P}{m} + V(X)$ (*i.e.*, friction included), the differential equation obtained in the second step would have *been a different differential equation*, so one would not be able to say that $\psi^*_{E,x}(x)$ is also an eigenfunction of H with eigenvalue E.

We shall see below that, for a bound-state energy E, the fact that, $\psi_{E,x}(x)$ and $\psi^*_{E,x}(x)$ are both eigenfunctions of H does not violate the previous no-degeneracy theorem.

With the above requirement on H, we may define two real linear combinations:

$$\psi_{E,x,r}(x) = \frac{1}{2}\left(\psi_{E,x}(x) + \psi^*_{E,x}(x)\right) \qquad \psi_{E,x,i}(x) = \frac{1}{2\,i}\left(\psi_{E,x}(x) - \psi^*_{E,x}(x)\right)$$

Because they are linear combinations of eigenstates with the same eigenvalue E, these two states are also eigenstates of H with eigenvalue E. Note that $\psi_{E,x,i}(x)$ is not just the imaginary part of $\psi_{E,x}(x)$; the extra factor of $1/i$ makes it real also. Hence, the theorem is proven – we constructed two real eigenfunctions from the two eigenfunctions $\psi_{E,x}(x)$ and $\psi^*_{E,x}(x)$.

Theorems on One-Dimensional States (cont.)

Now the issue with degeneracy. If E is a free-state energy, then there is no problem with having two such independent functions; there is no guarantee against degeneracy for free states.

What happens when E is a bound-state energy? Do we still get two real solutions? For E a bound-state energy, the no-degeneracy theorem implies $\psi^*_{E,x}(x) = \alpha\,\psi_{E,x}(x)$. Moreover, since we know $|\psi^*_{E,x}(x)|^2 = |\psi_{E,x}(x)|^2$, we have $|\alpha|^2 = 1$ and thus $\alpha = e^{i\theta}$. Therefore, we find

$$\psi_{E,x,r}(x) = \frac{1}{2}\left(1 + e^{i\theta}\right)\psi_{E,x}(x) \qquad \psi_{E,x,i}(x) = \frac{1}{2i}\left(1 - e^{i\theta}\right)\psi_{E,x}(x)$$

Thus, we get two real functions that are just $\psi_{E,x}(x)$ multiplied by two different complex numbers; they are the same up to a multiplicative factor and thus not linearly independent. So the no-degeneracy theorem is not violated. In special cases, in fact, one or the other vanishes identically: If one starts out with $\psi_{E,x}(x)$ perfectly real, then $e^{i\theta} = 1$ and $\psi_{E,x,i}(x) = 0$; if one starts with $\psi_{E,x}(x)$ perfectly imaginary, then $e^{i\theta} = -1$ and $\psi_{E,x,r}(x) = 0$. It all holds together.

Section 6
The One-Dimensional Simple Harmonic Oscillator

Lecture 21:
Simple Harmonic Oscillator: Coordinate Basis

Date Given: 2008/11/17
Date Revised: 2008/11/17

Motivation

As you have no doubt heard before, the primary motivation for studying the simple harmonic oscillator is that, for any system subject to a potential energy $V(x)$ and for motion around an equilibrium position x_0 (where, by definition, $\frac{d}{dx} V(x)\Big|_{x_0} = 0$), the system acts like a simple harmonic oscillator. Explicitly, the potential energy is

$$V(x) = V(x_0) + \frac{d}{dx} V(x)\Big|_{x_0} (x - x_0) + \frac{1}{2} \frac{d^2}{dx^2} V(x)\Big|_{x_0} (x - x_0)^2 + \cdots$$

The first term is an unimportant constant, the second term vanishes at x_0 because it is an equilibrium position, so the third term is the first important term. It is quadratic in the displacement from x_0, just like a simple harmonic oscillator. If the kinetic term is the usual $\frac{p^2}{2\,m}$, then the Hamiltonian for the system may be approximated as

$$\mathcal{H}(x, p) = \frac{p^2}{2\,m} + \frac{1}{2}\,k\,x^2$$

where we define $k = \frac{d^2}{dx^2} V(x)\Big|_{x_0}$ and redefine the origin to be at x_0. That is, we have the simple harmonic oscillator.

Motivation (cont.)

The above argument is equally valid for multiparticle systems; in fact, the SHO approximation can be even more useful there because of the complication of dealing with so many particles.

See Shankar and any intermediate mechanics textbook (Marion and Thornton, Hand and Finch, Goldstein) for more examples.

$\{|x\rangle\}$-Basis Hamiltonian and Eigenvalue-Eigenvector Equation

The classical and quantum Hamiltonians are

$$\mathcal{H}(x,p) = \frac{p^2}{2\,m} + \frac{1}{2}\,m\,\omega^2\,x^2 \qquad \Longrightarrow \qquad H(X,P) = \frac{P^2}{2\,m} + \frac{1}{2}\,m\,\omega^2\,X^2 \tag{6.1}$$

As usual, we first need to find the eigenstates of the Hamiltonian, $H\,|\psi_E\,\rangle = E\,|\psi_E\,\rangle$. Also as usual, we insert completeness and take the product on the left with $\langle x\,|$ (repeating this completely generic step so that you are reminded of it!):

$$\langle x\,|H\,|\psi_E\,\rangle = \langle x\,|E\,|\psi_E\,\rangle$$

$$\int_{-\infty}^{\infty} dx'\,\langle x\,|H\,|x'\,\rangle\langle x'\,|\psi_E\,\rangle = E\,\langle x\,|\psi_E\,\rangle$$

$$\int_{-\infty}^{\infty} dx'\,\langle x\,|\left(\frac{P^2}{2\,m} + \frac{1}{2}\,m\,\omega^2\,X^2\right)|x'\,\rangle\,\psi_{E,x}(x') = E\,\psi_{E,x}(x)$$

Coordinate Basis (cont.)

We calculated in Equations 5.27 and 5.28 the matrix elements $\langle x\,|P^2\,|x'\,\rangle$ and $\langle x\,|V(X)\,|x'\,\rangle$ in a way valid for any one-dimensional problem, so we use those results:

$$\int_{-\infty}^{\infty} dx\,\delta(x-x')\left[-\frac{\hbar^2}{2\,m}\,\frac{d}{d(x')^2}+\frac{1}{2}\,m\,\omega^2\,(x')^2\right]\psi_{E,x}(x')=E\,\psi_{E,x}(x)$$

$$-\frac{\hbar^2}{2\,m}\,\frac{d}{dx^2}\,\psi_{E,x}(x)+\frac{1}{2}\,m\,\omega^2\,x^2\,\psi_{E,x}(x)=E\,\psi_{E,x}(x) \qquad (6.2)$$

There are two natural scales, a length $b=\sqrt{\frac{\hbar}{m\,\omega}}$ and an energy $E_0=\hbar\,\omega=\frac{\hbar^2}{m\,b^2}$. If we define $y=\frac{x}{b}$ and $\varepsilon=\frac{E}{E_0}$, then we may scale out the dimensions:

$$b^2\,\frac{d}{dx^2}\,\psi_{E,x}(x)+\frac{2\,m\,E\,b^2}{\hbar^2}\,\psi_{E,x}(x)-\frac{m^2\omega^2\,b^4}{\hbar^2}\,\frac{x^2}{b^2}\,\psi_{E,x}(x)=0$$

$$\frac{d^2}{dy^2}\,\psi_\varepsilon(y)+\left(2\,\varepsilon-y^2\right)\psi_\varepsilon(y)=0 \qquad (6.3)$$

The physics is now mostly done and we have a math problem. The other bit of input from the physics is the boundary condition $\psi_\varepsilon(y)\to 0$ faster than $1/\sqrt{|y|}$ as $|y|\to\infty$ so that the resulting state can be normalized. We know the state should be normalizable to unity because the potential becomes infinite at large $|x|$ and thus there can be no free states.

Solving the Differential Equation

We must now find the solutions of Equation ??

$$\frac{d^2}{dy^2}\,\psi_\varepsilon(y) + \left(2\,\varepsilon - y^2\right)\psi_\varepsilon(y) = 0$$

subject to the boundary condition $\psi_\varepsilon(y) \to 0$ faster than $1/\sqrt{|y|}$ as $|y| \to \infty$. ε is a parameter that gives the energy eigenvalue.

This is a second-order linear differential equation with non-constant, polynomial coefficients, so you know from your math classes that one has to construct a series solution. In principle, this is all straightforward. However, we will go through much of the calculation in detail because this is the first such case we have encountered. Also, for the sake of your mathematical physics education, it is important to become adept at doing this kind of thing: just as differentiation and integration are second nature to you by now, and hopefully you are getting to a similar point on linear algebra, you need to internalize methods of solving differential equations.

Coordinate Basis (cont.)

Let's first consider the asymptotic behavior of the equation: this will put the problem in a cleaner form for the series solution. At large $|y|$, $y^2 \gg \varepsilon$, so we have

$$\frac{d^2}{dy^2}\psi_\varepsilon(y) - y^2\psi_\varepsilon(y) = 0$$

(Notice that the equation no longer depends on ε and hence we can drop the ε label for now.) A solution to this equation, *in the same limit* $|y| \to \infty$, is

$$\psi(y) = A\, y^m e^{\pm\frac{y^2}{2}}$$

which we can see by direct substitution:

$$\begin{aligned}
\frac{d^2}{dy^2}\psi(y) &= \frac{d}{dy}\left(\left[m\, y^{m-1} + y^m\,(\pm y)\right] A\, e^{\pm\frac{y^2}{2}}\right) \\
&= \left[\left(m\,(m-1)\, y^{m-2} \pm (m+1)\, y^m\right) + \left(m\, y^{m-1} \pm y^{m+1}\right)(\pm y)\right] A\, e^{\pm\frac{y^2}{2}} \\
&= \left[y^{m+2} \pm (2\,m+1)\; y^m + m\,(m-1)\, y^{m-2}\right] A\, e^{\pm\frac{y^2}{2}}
\end{aligned}$$

Coordinate Basis (cont.)

Now take the $|y| \to \infty$ limit:

$$\frac{d^2}{dy^2}\psi(y) \overset{|y|\to\infty}{\longrightarrow} y^2\, A\, y^m e^{\pm\frac{y^2}{2}} = y^2\psi(y)$$

The asymptotic solution works. So, our solution must asymptote to $y^m\, e^{\pm\frac{y^2}{2}}$ at large $|y|$. To be normalizable, and hence physically allowed, the $+$ solution is disallowed, leaving only

$$\psi(y) \overset{|y|\to\infty}{\sim} y^m\ e^{-\frac{y^2}{2}} \tag{6.4}$$

Coordinate Basis (cont.)

In order to further constrain the form of the solution, let us consider the $|y| \ll \varepsilon$ limit, in which we can ignore the y^2 term, giving

$$\frac{d^2}{dy^2}\,\psi_\varepsilon(y) + 2\,\varepsilon\,\psi_\varepsilon(y) = 0$$

This is a second-order linear differential equation with constant coefficients, so we know the solution is a sum of harmonic functions:

$$\psi_\varepsilon(y) = \alpha\,\cos\left(\sqrt{2\,\varepsilon}\,y\right) + \beta\,\sin\left(\sqrt{2\,\varepsilon}\,y\right)$$

Notice that the solution depends on ε in this limit. Since we ignored the term of order y^2 in the differential equation in this limit, we only need to consider this solution to first order in y, which gives

$$\psi_\varepsilon(y) \overset{|y|\to 0}{\longrightarrow} \alpha + \beta\,\sqrt{2\,\varepsilon}\,y \tag{6.5}$$

That is, the solution behaves like a polynomial as $|y| \to 0$.

Coordinate Basis (cont.)

A full solution that would satisfy the above limits is

$$\psi_\varepsilon(y) = u_\varepsilon(y)\, e^{-\frac{y^2}{2}} \tag{6.6}$$

where $u(y) \overset{|y|\to 0}{\longrightarrow} \alpha + \beta\sqrt{2\,\varepsilon}\,y$ and $u(y) \overset{|y|\to\infty}{\longrightarrow} y^m$. (Note that $u_\varepsilon(y)$ now carries the ε subscript because the Gaussian portion has no dependence on ε by construction.) Let us plug this into the full differential equation and obtain a differential equation for $u(y)$:

$$\left[\frac{d^2}{dy^2} + (2\,\varepsilon - y^2)\right]\left(u_\varepsilon(y)\, e^{-\frac{y^2}{2}}\right) = 0$$

$$\left[\frac{d^2}{dy^2}\,u_\varepsilon(y) - 2\,y\,\frac{d}{dy}\,u_\varepsilon(y) + \left(y^2 - 1 + 2\,\varepsilon - y^2\right) u_\varepsilon(y)\right] e^{-\frac{y^2}{2}} = 0$$

$$\frac{d^2}{dy^2}\,u_\varepsilon(y) - 2\,y\,\frac{d}{dy}\,u_\varepsilon(y) + (2\,\varepsilon - 1)\,u_\varepsilon(y) = 0 \tag{6.7}$$

Our asymptotic considerations indicate that the solution to this differential equation behaves like a polynomial both as $|y| \to 0$ and as $|y| \to \infty$. This leads us to try a series solution of the form $u_\varepsilon(y) = \sum_{n=0}^{\infty} C_{\varepsilon,n} y^n$.

Coordinate Basis (cont.)

Feeding the series solution form into the differential equation yields

$$\sum_{n=0}^{\infty} C_{\varepsilon,n}\left[n\,(n-1)\,y^{n-2} - 2\,n\,y^{n} + (2\,\varepsilon - 1)\,y^{n}\right] = 0$$

$$\sum_{m=0}^{\infty} C_{\varepsilon,m+2}\,(m+2)\,(m+1)\,y^{m} = \sum_{n=0}^{\infty} C_{\varepsilon,n}\,(2\,n+1-2\,\varepsilon)\,y^{n}$$

where, for the first piece, we relabeled the sum over n to be a sum over $m = n-2$; the m index starts at 0, not -2, because the first two terms of that series vanish ($n\,(n-1) = 0$ for $n = 0, 1$), and we moved the second and third pieces to the right side. Since the functions $\{y^n\}$ are linearly independent (recall, we argued that they could be used as a basis for a function space on the interval $[a, b]$ because no one of them can be written as sum of the others), the sums must be equal term-by-term, so

$$C_{\varepsilon,n+2} = C_{\varepsilon,n}\frac{2\,n+1-2\,\varepsilon}{(n+2)(n+1)} \tag{6.8}$$

The coefficients $C_{\varepsilon,0}$ and $C_{\varepsilon,1}$ are left to be determined by initial conditions.

Coordinate Basis (cont.)

Now, as is usual in these circumstances, we require that the series terminate at some point so that the asymptotic behavior is obeyed: as we have it now, the series goes on forever rather than converging to y^m for some m. This also explains why we have obtained no quantization condition on ε yet: as we explained in Section 5.3, quantization of the energy arises because of the bound-state condition that the solution must decay sufficiently quickly at ∞; our solution does not yet satisfy this condition!

Shankar complicates this issue by following the unterminated solution to its logical conclusion. That is unnecessary: we know that $\sum_{n=0}^{\infty} C_n y^n \overset{|y|\to\infty}{\longrightarrow} y^m$ is impossible unless $C_n = 0$ for $n > m$ by the same linear independence argument as made before.

Coordinate Basis (cont.)

So, let's just require termination:

$$0 = C_{\varepsilon,n+2} = C_{\varepsilon,n}\frac{2n+1-2\varepsilon}{(n+2)(n+1)}$$
$$0 = 2n+1-2\varepsilon$$
$$\varepsilon = \frac{2n+1}{2}$$

We obtain the condition that ε must be an odd half-integer.

The above condition only terminates either the odd or even coefficients, depending on whether the n set by ε is odd or even. To ensure termination, we must require $C_{\varepsilon,1} = 0$ when n is even so that the odd terms all vanish and, conversely, $C_{\varepsilon,0} = 0$ when n is odd so all the even powers vanish.

To summarize, with ε an odd half-integer of the form $\varepsilon = \frac{2n+1}{2}$, $u_\varepsilon(y)$ is a polynomial of order $n = \frac{2\varepsilon-1}{2}$ and containing only the odd or even powers of y depending on whether n is even or odd, respectively. This solution matches our asymptotic conditions.

Explicit Form for the SHO Solutions

Our solutions are of the form

$$\psi_{E_n,x}(x) = \left(\frac{m\,\omega}{\pi\,\hbar\,2^{2n}\,(n\,!)^2}\right)^{1/4} H_n\left(x\sqrt{\frac{m\,\omega}{\hbar}}\right) e^{-\frac{m\,\omega\,x^2}{2\,\hbar}} \qquad E_n = \left(n+\frac{1}{2}\right)\hbar\omega \tag{6.9}$$

where $H_n(y)$ are the Hermite polynomials

$$H_0(y) = 1 \qquad H_1(y) = 2\,y$$

$$H_2(y) = -2\left(1-2\,y^2\right) \qquad H_3(y) = -12\left(y-\frac{2}{3}\,y^3\right)$$

where the choice of the C_0 and C_1 coefficients in each case is arbitrary but is a convention (one that allows the given simple form for the normalization). The normalization can be calculated by some tedious integrations that we will not go through. The different eigenfunctions are of course orthonormal, $\langle\psi_{E_n}\,|\psi_{E_m}\,\rangle = \int_{-\infty}^{\infty} dx\,\psi^*_{E_n,x}(x)\,\psi_{E_m,x}(x) = \delta_{nm}$. The related orthogonality and normalization of the Hermite polynomials alone is given in Shankar, as well as a recurrence relation that we will not need.

Lecture 22:
Simple Harmonic Oscillator: Energy Basis

Date Given: 2008/11/19
Date Revised: 2008/11/19

Properties of the SHO Solutions

We note some interesting properties of the solutions:

- **Parity**
  The parity operator Π makes the transformation on the wavefunction $x \longrightarrow -x$; formally, $\Pi|x\rangle = |-x\rangle$. Since this defines the action of Π for an orthonormal basis, it fully defines Π. We can see Π is Hermitian:

$$\langle x'|\Pi|x\rangle = \langle x'|-x\rangle = \delta(x'+x) = \langle -x'|x\rangle \quad \Longrightarrow \quad \langle x'|\Pi = \langle -x'|$$

  Hence, it has real eigenvalues and its eigenvectors form a complete basis. The symmetry of the potential $V(x)$ implies that the Hamiltonian and the parity operator commute, so the eigenvectors of H must also be eigenvectors of Π. Since Π is Hermitian and $\Pi^2 = I$, the allowed eigenvalues of Π are ± 1, corresponding to even and odd functions, respectively. We see this property reflected in the SHO eigenfunctions. As we explained in deriving them, each Hermite polynomial contains only even or odd powers of the argument. Hence, they are either even or odd in their argument. The Gaussian that multiplies them is even, so the solutions themselves are either even or odd in x. The same property held or the eigenfunctions of the particle in a box, and in fact is the reason we chose to have the box over the interval $\left[-\frac{L}{2}, \frac{L}{2}\right]$ rather than $[0, L]$.

- **Number and Positions of Zeroes**
  The Hermite polynomial H_n is a polynomial of order n, so it must have n zeros. What is interesting is that these zeroes are contained inside the classical turning points, $x_0 = \pm\sqrt{\frac{2\,E_n}{k}}$, the points where the kinetic energy vanishes and the potential energy is maximized. The argument is not trivial and goes as follows. We may calculate the first derivative at any point by by

$$\frac{d}{dy}\,\psi_\varepsilon(y')\Big|_y^\infty = \int_y^\infty dy'\,\frac{d^2}{d(y')^2}\,\psi_\varepsilon(y') = \int_y^\infty dy'\,\left((y')^2 - 2\,\varepsilon\right)\psi_\varepsilon(y')$$

Since $\psi_\varepsilon(y) \overset{|y|\to\infty}{\longrightarrow} 0$ in order for $\psi_\varepsilon(y)$ to be normalizable, it holds that $\frac{d}{dy}\,\psi_\varepsilon(y) \overset{|y|\to\infty}{\longrightarrow} 0$ also. Thus, one end of the left side vanishes, giving

$$\frac{d}{dy}\,\psi_\varepsilon(y) = -\int_y^\infty dy'\,\left[(y')^2 - 2\,\varepsilon\right]\psi_\varepsilon(y')$$

Note that the factor in the integrand, $(y')^2 - 2\,\varepsilon$, is positive for $(y')^2 > y_0^2 = 2\,\varepsilon$.

Coordinate Basis (cont.)

Next, define y_n, with $\varepsilon = n + \frac{1}{2}$, to be the position of the last zero of $\psi_\varepsilon(y)$ (there is no amibiguity as to what y_0 means because we know the $n = 0$ mode has no zeros). For $y > y_n$, the sign of $\psi_\varepsilon(y)$ is fixed because there are no more zeros. For specificity, suppose it is positive; we can always apply a -1 to $\psi_\varepsilon(y)$ to make this true without changing the positions of the zeros. The above equation tells us the first derivative must be negative for $y \geq \max(y_0, y_n)$ because the integrand is positive in this region.

Now, suppose $y_n \geq y_0$. Since $\psi_\varepsilon(y)$ is positive for the regime $y > y_n$, the derivative at y_n, $\frac{d}{dy}\psi_\varepsilon(y_n)$, must be positive in order for the function to cross through zero from negative to positive value at y_n. But we showed above that the first derivative is negative for $y \geq \max(y_0, y_n)$, which, by our assumption $y_n \geq y_0$, corresponds to $y \geq y_n$. We have a contradiction.

The contradiction can be resolved by supposing instead $y_0 > y_n$. Then, the integrand can go through zero to negative values for $y > y_n$, making it possible for the integral to change sign and for $\frac{d}{dy}\psi_\varepsilon(y_n)$ to go from its asymptotic negative value to a positive value at y_n.

So we have $y_n < y_0 = \sqrt{2\,\varepsilon}$. Putting the units back in, we have

$$x_n < y_0\sqrt{\frac{\hbar}{m\,\omega}} = \sqrt{\frac{\hbar}{m\,\omega}}\sqrt{2\left(n+\frac{1}{2}\right)} = \sqrt{\frac{2\,E_n}{k}} = x_0$$

Coordinate Basis (cont.)

The same argument can be carried through for negative x; QED.

One last note. The above implies that the wavefunctions are oscillatory only in the region inside the classical turning points, and decaying outside there. This is consistent with our generic discussion of bound states and how quantization arises in Section 5.3. There, by breaking the potential into an infinite number of infinitesimally wide piecewise constant regions, we saw that solutions ought to be oscillatory in regions where $E > V(x)$ and decay outside there. The SHO solutions obey this rule.

- **Position and Momentum Uncertainties**
  Shankar goes through a straightforward demonstration that the state that saturates the uncertainty principle, with $\sqrt{\langle(\Delta X)^2\rangle\langle(\Delta P)^2\rangle} = \frac{\hbar}{2}$, is the ground state of the SHO. This is unusual; we found for the particle in a box, that the energy of the ground state is higher than that implied by the uncertainty principle.

  Shankar also discusses how this implies there is zero-point energy — just as for the particle in a box, the ground state has nonzero energy — and that this has measurable physical consequences, such as the fact that the vibrational energy content of a crystal at absolute zero does not vanish.

Coordinate Basis (cont.)

- **Classical Limit**
  Shankar also shows by illustration for $n = 11$ that the higher n modes have probability disributions peaked at the turning points, corresponding to the distribution of dwell time one expects for a classical SHO, in which more time is spent near the turning points than near the origin. The lowest modes do not satisfy this, but we only expect to recover the classical dwell time in the limit of high energy, when the discretization of the allowed energies becomes unnoticeable; *i.e.*, when $E \gg \hbar\omega$.

The Propagator

We calculate the propagator by the standard formula, Equation 4.18, and also calculate its $\{|x\rangle\}$-basis matrix elements:

$$U(t) = e^{-\frac{i}{\hbar}Ht} = \sum_n e^{-\frac{i}{\hbar}E_n t}|\psi_{E_n}(0)\rangle\langle\psi_{E_n}(0)| = \sum_{n=0}^{\infty} e^{-i\left(n+\frac{1}{2}\right)\omega t}|\psi_{E_n}\rangle\langle\psi_{E_n}|$$

$$\begin{aligned}[U(t)]_{xx'} &= \langle x|U(t)|x'\rangle = \sum_{n=0}^{\infty} e^{-i\left(n+\frac{1}{2}\right)\omega t}\langle x|\psi_{E_n}\rangle\langle\psi_{E_n}|x'\rangle \\ &= \sum_{n=0}^{\infty} e^{-i\left(n+\frac{1}{2}\right)\omega t}\left(\frac{m\omega}{\pi\hbar 2^{2n}(n!)^2}\right)^{1/2} H_n(x)\, e^{-\frac{m\omega x^2}{2\hbar}} H_n(x')\, e^{-\frac{m\omega (x')^2}{2\hbar}} \\ &= \left(\frac{m\omega}{2\pi i\hbar \sin\omega t}\right)^{1/2} \exp\left[\frac{i}{\hbar}\, m\omega \frac{\left(x^2+(x')^2\right)\cos\omega t - 2xx'}{2\sin\omega t}\right]\end{aligned}$$

where the sum is not obvious. Shankar derives it using path integrals, which we will unfortunately not have time to cover.

Motivation

So far, we have found the SHO's energy eigenvalues and eigenstates in the $\{|x\rangle\}$ basis. However, because the energies go up linearly with the state index n (*i.e.*, the energy levels are spaced equally), we are motivated to think of the eigenstate that the system is in as corresponding to a number of "quanta" of energy $\hbar\omega$. Rather than thinking in terms of energy levels, we could think in terms of a "number" operator multiplying the energy per quantum. How do we get there?

Before continuing, we remind the reader that it is generically true that the Hamiltonian can be written in the form

$$H = \sum_n E_n |\psi_{E_n}\rangle\langle\psi_{E_n}| \equiv \sum_n E_n |E_n\rangle\langle E_n| \equiv \sum_n E_n |n\rangle\langle n|$$

The issue at hand is whether this offers any simplification. In general, it does not because, in general, the E_n are not equally spaced. In fact, for potentials that have both bound states and free states, there are two pieces to the sum involved, one for the discretely spaced bound states and one for the continuum of free states. There is no natural idea of an energy "quantum" and hence no motivation to think in terms of a number operator.

Energy Basis (cont.)

Raising and Lowering Operators

Without real motivation, we define the operator

$$a = \sqrt{\frac{1}{2}\frac{m\omega}{\hbar}}\,X + i\sqrt{\frac{1}{2}\frac{1}{m\omega\hbar}}\,P$$

Its adjoint is

$$a^\dagger = \sqrt{\frac{1}{2}\frac{m\omega}{\hbar}}\,X - i\sqrt{\frac{1}{2}\frac{1}{m\omega\hbar}}\,P$$

(After you have seen canonical transformations in Ph106, the motivation will be clear; we will come back to this later.) One can easily check that they satisfy

$$\left[a, a^\dagger\right] = 1 \qquad H = \left(a^\dagger a + \frac{1}{2}\right)\hbar\omega \quad \Longleftrightarrow \quad \widehat{H} = \frac{H}{\hbar\omega} = \left(a^\dagger a + \frac{1}{2}\right)$$

Both relations arise from $[X, P] = i\,\hbar$, which holds in any basis.

We call $a^\dagger$ and a raising and lowering operators for reasons that will become apparent.

Basis-Free Eigenstates of the Hamiltonian

We want to find the eigenstates of $\widehat{H}$. To date, we have always had to write the eigenvalue-eigenvector equation for the Hamiltonian in either the position or momentum basis; usually the former to obtain a differential equation that we can solve for the $\{|x\rangle\}$-basis representation of the eigenstates $|E\rangle$. Here, we are going to try to avoid a basis altogether. Getting used to this idea is not easy.

So, we want to solve

$$\widehat{H}\,|\varepsilon\rangle = \varepsilon\,|\varepsilon\rangle$$

(The use of the ε symbol is suggestive.) First, we need the following:

$$[a, \widehat{H}] = \left[a, a^\dagger a + \frac{1}{2}\right] = a\,a^\dagger a - a^\dagger a\,a = \Big([a, a^\dagger] + a^\dagger a\Big)\,a - a^\dagger a\,a = a$$

$$[a^\dagger, \widehat{H}] = \left[a^\dagger, a^\dagger a + \frac{1}{2}\right] = a^\dagger a^\dagger a - a^\dagger a\,a^\dagger = a^\dagger a^\dagger a - a^\dagger\Big([a, a^\dagger] + a^\dagger a\Big) = -a^\dagger$$

Energy Basis (cont.)

The above lets us show that, if $|\varepsilon\rangle$ is an eigenstate of $\widehat{H}$ with eigenvalue ε, then a and $a^\dagger$ generate from $|\varepsilon\rangle$ other eigenstates of $\widehat{H}$:

$$\widehat{H}\,a\,|\varepsilon\rangle = \left([\widehat{H},a] + a\,\widehat{H}\right)|\varepsilon\rangle = (-a + a\,\varepsilon)\,|\varepsilon\rangle = (\varepsilon - 1)\,a\,|\varepsilon\rangle$$
$$\widehat{H}\,a^\dagger|\varepsilon\rangle = \left([\widehat{H},a^\dagger] + a^\dagger\widehat{H}\right)|\varepsilon\rangle = \left(a^\dagger + a^\dagger\varepsilon\right)|\varepsilon\rangle = (\varepsilon + 1)\,a^\dagger|\varepsilon\rangle$$

That is

$$a\,|\varepsilon\rangle = C_\varepsilon|\varepsilon - 1\rangle \qquad a^\dagger|\varepsilon\rangle = C_{\varepsilon+1}|\varepsilon + 1\rangle$$

(We use the fact that there is no degeneracy of bound states in one dimension, so there is one eigenstate with eigenvalue $|\varepsilon\rangle$.) The rationale for calling $a^\dagger$ and a raising and lowering operators now becomes clear.

Energy Basis (cont.)

Left with the above only, we might think that there are states of arbitrarily low energy. We know from our $\{|x\rangle\}$-basis solution that this is not true. We can also prove it more generally given the form of the Hamiltonian. Consider $\langle\psi|H|\psi\rangle$ for an arbitrary state $|\psi\rangle$:

$$\langle\psi|H|\psi\rangle = \frac{1}{2m}\langle\psi|P^2|\psi\rangle + \frac{1}{2}k\langle\psi|X^2|\psi\rangle = \frac{1}{2m}|P|\psi\rangle|^2 + \frac{1}{2}k|X|\psi\rangle|^2 \geq 0$$

Therefore, any eigenvalues of H and hence $\widehat{H}$ must be nonnegative. Notice that the proof that the kinetic term is nonnegative holds always, but the proof for the potential term is specific to the SHO potential and is not generally true.

The lower limit tells us there must be a lowest energy state that satisfies

$$a|\varepsilon_0\rangle = 0$$

so that we cannot obtain any lower energy states. (We should properly write $|0\rangle$ on the right side for the null vector, but, as we shall see, this may become confused with $|0\rangle = |\varepsilon_0\rangle$, the ground state ket.) We then have

$$\langle\varepsilon_0|a^\dagger a|\varepsilon_0\rangle = 0 \implies \langle\varepsilon_0|\widehat{H}|\varepsilon_0\rangle = \frac{1}{2} \text{ and } \langle\varepsilon_0|H|\varepsilon_0\rangle = \frac{1}{2}\hbar\omega$$

Energy Basis (cont.)

Are we certain there is no parallel chain of states with different energies? Since we have started from scratch in the energy basis, we must assume we have no prior knowledge that ε must be an integer.

Yes, we can eliminate the possibility of a parallel chain by calculating the energy of the lowest energy state. Suppose there were a parallel chain of states, $|\varepsilon' \rangle$, with eigenvalues ε'. By the same argument as above, we are assured there is a lowest energy state in the chain, $|\varepsilon'_0 \rangle$ for which $a|\varepsilon'_0 \rangle = 0$. The dimensionless energy of that state is then also $\varepsilon'_0 = \frac{1}{2}$ by the same argument as for $|\varepsilon_0 \rangle$. So $|\varepsilon'_0 \rangle$ is degenerate with $|\varepsilon_0 \rangle$. But we have proven that one-dimensional bound states are nondegenerate, so it must be that $|\varepsilon'_0 \rangle$ is proportional to $|\varepsilon_0 \rangle$; they are the same state.

So, we have the set of eigenstates of H, connected by the raising and lowering operators.

The Inner Product in the Energy Basis

At this point, we only have made use of the following facts:

- X and P, and therefore H, are Hermitian operators on the linear vector space of states.
- Being Hermitian, we are assured that the eigenstates of H form a basis for the space.
- Because the potential has no maximum value, we are assured that all states are bound states and, with the addition of the fact that we are considering a one-dimensional potential, we are assured they are nondegenerate. (Though we proved nondegeneracy via the position basis, nondegeneracy is a non-basis-specific property; moreover, our use of the position basis was not specific to the SHO.)
- We can define raising and lowering operators $a^\dagger$ and a in terms of X and P, and we find they connect the various eigenstates.

Energy Basis (cont.)

We have not formally defined the inner product on this space. (The inner product we have defined for our one-dimensional problems so far relies on the position basis, which we are trying to avoid.) However, we proved in Section 3.6 that, if there is a reasonable inner product for the space (recall, Hermiticity does not require the existence of an inner product, though it is much easier to prove when there is one!), then the eigenstates of a Hermitian operator are mutually orthogonal. We therefore take this as part of the *definition* of the inner product on the space. This leaves the normalization of the inner product unspecified, so we also assume the eigenstates are normalized. Since the C_ε constants are still undetermined, we still have the necessary freedom to set the normalizations of the eigenstates individually.

Let us now use that freedom to determine what the constants C_ε must be in order to be consistent with the above normalization choice. We obtain a condition on C_ε as follows, using $a\,|\varepsilon\,\rangle = C_\epsilon|\varepsilon - 1\,\rangle$ from earlier:

$$\begin{aligned} 1 = \langle \varepsilon - 1\,|\varepsilon - 1\,\rangle &= |C_\varepsilon|^{-2}\langle \varepsilon\,|a^\dagger a\,|\varepsilon\,\rangle = |C_\varepsilon|^{-2}\langle \varepsilon\,|\left(\widehat{H} - \frac{1}{2}\right)|\varepsilon\,\rangle \\ &= |C_\varepsilon|^{-2}\left(\varepsilon - \frac{1}{2}\right)\langle \varepsilon\,|\varepsilon\,\rangle = |C_\varepsilon|^{-2}\left(\varepsilon - \frac{1}{2}\right) \\ \Longrightarrow \qquad |C_\varepsilon|^2 &= \varepsilon - \frac{1}{2} \end{aligned}$$

Energy Basis (cont.)

At this point, life is less confusing if we define the number operator, $N = a^\dagger\, a = \widehat{H} - \frac{1}{2}$. It obviously has the same eigenstates as $\widehat{H}$ and H, with eigenvalues $n = \varepsilon - \frac{1}{2}$. It "counts" the quanta of energy in the oscillator. The above relation is then

$$|C_n|^2 = n \qquad \Longleftrightarrow \qquad C_n = e^{\,i\,\phi_n}\sqrt{n}$$

where ϕ_n is arbitrary for each n. The simplest convention is $\phi_n = 0$ for all n, giving

$$a\,|n\,\rangle = \sqrt{n}\,|n-1\,\rangle \qquad a^\dagger|n\,\rangle = \sqrt{n+1}\,|n+1\,\rangle \qquad \langle n\,|m\,\rangle = \delta_{nm}$$

This defines the inner product for the entire space because the $\{|n\,\rangle\}$ are a basis. Any state can now be obtained from the ground state by use of raising operators:

$$|n\,\rangle = \frac{1}{\sqrt{n\,!}}\left(a^\dagger\right)^n|0\,\rangle$$

Finally, we make the important note that, given the above definition of $|n\,\rangle$, there is now a *notational degeneracy between the SHO ground state* $|0\,\rangle$ *and the null vector* $|0\,\rangle$. As we did above, we are simply going to use 0 for the null vector because context will always make it clear whether 0 refers to scalar or a vector.

Matrix Elements of Various Operators

Here we list some useful matrix elements that we can derive from what we have done so far:

$$\langle n\,|a\,|m\,\rangle = \sqrt{m}\,\langle n\,|m-1\,\rangle = \sqrt{m}\,\delta_{n,m-1}$$

$$\langle n\,|a^\dagger|m\,\rangle = \sqrt{m+1}\,\langle n\,|m+1\,\rangle = \sqrt{m+1}\,\delta_{n,m+1}$$

$$X = \sqrt{\frac{\hbar}{2\,m\,\omega}}\left(a^\dagger + a\right) \Longrightarrow \langle n\,|X\,|m\,\rangle = \sqrt{\frac{\hbar}{2\,m\,\omega}}\left(\sqrt{m+1}\,\delta_{n,m+1} + \sqrt{m}\,\delta_{n,m-1}\right)$$

$$P = i\sqrt{\frac{m\,\omega\,\hbar}{2}}\left(a^\dagger - a\right) \Longrightarrow \langle n\,|P\,|m\,\rangle = i\sqrt{\frac{m\,\omega\,\hbar}{2}}\left(\sqrt{m+1}\,\delta_{n,m+1} - \sqrt{m}\,\delta_{n,m-1}\right)$$

$$H = \left(N + \frac{1}{2}\right)\hbar\omega \Longrightarrow \langle n\,|H\,|m\,\rangle = \left(n + \frac{1}{2}\right)\hbar\omega\,\delta_{n,m}$$

The marix representations are given in Shankar; they are obvious from the above.

Energy Basis (cont.)

Shankar makes the point that working in the energy basis, which gives these simple forms for matrix elements of the fundamental X and P operators, makes matrix elements of *any* function of X and P (or of a and $a^\dagger$) easier to calculate than evaluating integrals using the $\{|x\rangle\}$-basis representation of the eigenstates. We won't try to reproduce the algebra of calculating $\langle 3|X^3|2\rangle$ here.

Lecture 23:
Simple Harmonic Oscillator:
Energy Basis – Coordinate Basis Correspondence
Rewriting Postulate 2

Date Given: 2008/11/21
Date Revised: 2008/11/21

From the Energy Basis to the Position Basis

We have made no reference to the $\{|x\rangle\}$ basis in our discussion of the energy basis so far, so we have no way to compute $P(x) = |\langle x|n\rangle|^2$, the probability density for finding a particle in state $|n\rangle$ at position x. There is a clever way to figure it out, though, using the raising and lowering operators. We begin with $a|0\rangle = 0$ (0 is the null vector here!) and take the inner product with $\langle x|$ and insert completeness:

$$
\begin{aligned}
0 = \langle x|a|0\rangle &= \int_{-\infty}^{\infty} dx' \, \langle x|a|x'\rangle\langle x'|0\rangle \\
&= \int_{-\infty}^{\infty} dx' \, \langle x| \left(\sqrt{\frac{m\omega}{2\hbar}} X + i \sqrt{\frac{1}{2m\omega\hbar}} P \right) |x'\rangle \, \psi_{0,x}(x') \\
&= \sqrt{\frac{m\omega}{2\hbar}} \int_{-\infty}^{\infty} dx' \left(x\,\delta(x-x') + \frac{\hbar}{m\omega} \left[\frac{d}{dx}\delta(x-x') \right] \right) \psi_{0,x}(x') \\
&= \sqrt{\frac{m\omega}{2\hbar}} \int_{-\infty}^{\infty} dx' \left(x\,\delta(x-x') + \frac{\hbar}{m\omega}\,\delta(x-x')\,\frac{d}{dx'} \right) \psi_{0,x}(x') \\
&= \sqrt{\frac{m\omega}{2\hbar}} \left(x + \frac{\hbar}{m\omega}\,\frac{d}{dx} \right) \psi_{0,x}(x)
\end{aligned}
$$

Energy Basis – Coordinate Basis Correspondence (cont.)

We thus have a first-order differential equation for $\psi_{0,x}(x)$:

$$\frac{1}{\sqrt{2}}\left(\sqrt{\frac{m\,\omega}{\hbar}}x+\sqrt{\frac{\hbar}{m\,\omega}}\,\frac{d}{dx}\right)\psi_{0,x}(x)=0\Longleftrightarrow\frac{1}{\sqrt{2}}\left(y+\frac{d}{dy}\right)\psi_0(y)=0$$

where, as before, we have switched to a dimensionless position variable, $y=\sqrt{\frac{m\,\omega}{\hbar}}\,x$. This is a first order differential equation, easily solved:

$$\frac{d\psi_0(y)}{\psi_0(y)}=-y\,dy$$

$$\ln\psi_0(y)+c=-\frac{y^2}{2}$$

$$\psi_0(y)=A\,e^{-\frac{y^2}{2}}$$

We thus recover the $\{|x\rangle\}$-basis representation of the ground state!

Energy Basis – Coordinate Basis Correspondence (cont.)

The appropriate normalization is $A = (m\omega/\pi\hbar)^{1/4}$. We may then use the raising operator to obtain the rest of the states:

$$\begin{aligned}
\langle x\,|n\rangle &= \langle x\,|\frac{\left(a^\dagger\right)^n}{\sqrt{n!}}\,|0\rangle \\
&= \frac{1}{\sqrt{n!}}\left(\sqrt{\frac{m\omega}{2\hbar}}\,X - i\sqrt{\frac{1}{2m\omega\hbar}}\,P\right)^n \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} \\
&= \frac{1}{\sqrt{n!}}\left(\sqrt{\frac{m\omega}{2\hbar}}\,x - \hbar\sqrt{\frac{1}{2m\omega\hbar}}\,\frac{d}{dx}\right)^n \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} \\
&= \left(\frac{m\omega}{\pi\hbar 2^{2n}(n!)^2}\right)^{1/4} \left(y - \frac{d}{dy}\right)^n e^{-\frac{y^2}{2}}
\end{aligned}$$

where we skipped the usual rigamarole of inserting completeness in the $\{|x\rangle\}$ basis between the $\left(a^\dagger\right)^n$ operator and $|0\rangle$; it goes the same way it did for our calculation of $\langle x\,|a\,|0\rangle$.

Energy Basis – Coordinate Basis Correspondence (cont.)

One can see that the above is equivalent to the formula for the $\{|x\rangle\}$-basis eigenstates (Equation 6.9) by using recursion relations for the Hermite polynomials. The two recursion relations given in Shankar (they are derived from the differential equation Equation 6.7 and the coefficient recursion relation Equation 6.8 for the Hermite polynomials) are

$$\frac{d}{dy}H_n(y) = 2\,n\,H_{n-1}(y) \qquad H_{n+1}(y) = 2\,y\,H_n(y) - 2\,n\,H_{n-1}(y)$$

which can be combined to find

$$H_{n+1}(y) = 2\,y\,H_n(y) - \frac{d}{dy}H_n(y)$$

Let's do the proof of equivalence inductively. First, show it is true for $n = 1$:

$$\begin{aligned}\langle x|1\rangle &= \left(\frac{m\,\omega}{\pi\,\hbar 2^2\,(1\,!)^2}\right)^{1/4}\left(y - \frac{d}{dy}\right)e^{-\frac{y^2}{2}} = \left(\frac{m\,\omega}{\pi\,\hbar 2^2\,(1\,!)^2}\right)^{1/4}(2y)\,e^{-\frac{y^2}{2}} \\ &= \left(\frac{m\,\omega}{\pi\,\hbar 2^2\,(1\,!)^2}\right)^{1/4}H_1(y)\,e^{-\frac{y^2}{2}} = \psi_{E_1,x}(y)\end{aligned}$$

as expected.

Next, assume it is true for n and show that it is true for $n+1$:

$$\begin{aligned}
\langle x\,|n+1\,\rangle &= \left(\frac{m\,\omega}{\pi\,\hbar\,2^{2(n+1)}\,[(n+1)\,!]^2}\right)^{1/4}\left(y-\frac{d}{dy}\right)^{n+1}e^{-\frac{y^2}{2}}\\
&= \left(\frac{1}{2^2(n+1)^2}\right)^{1/4}\left(y-\frac{d}{dy}\right)\left(\frac{m\,\omega}{\pi\,\hbar\,2^{2n}\,[n!]^2}\right)^{1/4}\left(y-\frac{d}{dy}\right)^{n}e^{-\frac{y^2}{2}}\\
&= \left(\frac{1}{2^2(n+1)^2}\right)^{1/4}\left(y-\frac{d}{dy}\right)\left(\frac{m\,\omega}{\pi\,\hbar\,2^{2n}\,[n!]^2}\right)^{1/4}H_n(y)e^{-\frac{y^2}{2}}\\
&= \left(\frac{m\,\omega}{\pi\,\hbar\,2^{2(n+1)}\,[(n+1)\,!]^2}\right)^{1/4}\left(2\,y\,H_n(y)-\frac{d}{dy}\,H_n(y)\right)e^{-\frac{y^2}{2}}\\
&= \left(\frac{m\,\omega}{\pi\,\hbar\,2^{2(n+1)}\,[(n+1)\,!]^2}\right)^{1/4}H_{n+1}(y)\,e^{-\frac{y^2}{2}} = \psi_{E_{n+1},x}(y)
\end{aligned}$$

We have thus found complete equivalence between the energy-basis eigenstates, reexpressed in the $\{|x\,\rangle\}$ basis, and our original $\{|x\,\rangle\}$-basis eigenstates.

Rewriting Postulate 2

We see from the energy-basis study of the SHO that the matrix elements of X and P in the $\{|x\rangle\}$ basis, which are specified by Postulate 2, were irrelevant. Recall Postulate 2:

The independent variables x and p that describe completely the state of a particle in classical mechanics are represented by Hermitian operators X and P in the Hilbert space of states, with X and P having the following matrix elements when using the position basis for the Hilbert space:

$$\langle x|X|x'\rangle = x\delta(x-x') \qquad \langle x|P|x'\rangle = -i\hbar\frac{d}{dx}\delta(x-x')$$

Any arbitrary classical dynamical variable $\omega(x,p)$ has a corresponding Hermitian operator

$$\Omega(X,P) = \omega(x\to X, p\to P)$$

where we simply replace x and p in ω with X and P to obtain $\Omega(X,P)$.

The part of the above statement that we used was the resulting commutation relation $[X,P] = i\hbar I$, along with the unintuitive but otherwise unsurprising definition of a.

Rewriting Postulate 2 (cont.)

We are thus led to a perhaps more fundamental statement of this postulate:

The independent variables x and p that describe completely the state of a particle in classical mechanics are represented by Hermitian operators X and P in the Hilbert space of states; X and P and are defined by the *canonical commutator*

$$[X, P] = i\hbar I$$

Any arbitrary classical dynamical variable $\omega(x, p)$ has a corresponding Hermitian operator

$$\Omega(X, P) = \omega(x \to X, p \to P)$$

where we simply replace x and p in ω with X and P to obtain $\Omega(X, P)$.

Rewriting Postulate 2 (cont.)

The first obvious question is: how do you figure out what the matrix elements of X and P are? Our work in the energy basis showed that you do not necessarily need to know off the bat what these are; they can be inferred from solving for the eigenstates in the energy basis.

The second obvious question is: is this consistent with our original version of Postulate 2? Let's show that this prescription works out in the X basis, meaning that it returns our original version of Postulate 2. We work in the basis of the eigenstates of X; we are guaranteed such a basis exists because X is Hermitian. For obvious reasons, we denote these basis states by $\{|x\rangle\}$. In this basis, it by definition holds that

$$X|x\rangle = x|x\rangle$$

If we normalize $\{|x\rangle\}$ such that $\langle x|x'\rangle = \delta(x - x')$, then we have

$$\langle x|X|x'\rangle = x\,\delta(x - x')$$

Rewriting Postulate 2 (cont.)

We know from Section 3.10 that, if we we choose $P = -i\,\hbar\,D$, where D is the derivative operator, then we obtain

$$\langle x\,|P|x'\,\rangle = -i\,\hbar\,\frac{d}{dx}\,\delta(x - x')$$

and that this leads to Equation 3.151,

$$[X, P] = i\,\hbar$$

The question is: is this the only possible valid definition of P? No. Since we are calculating the commutator with X, we can add to the above definition any operator that commutes with X. Since the only operators we know about are X and P (every other operator we have looked at has been derived from them), the only kind of operator that commutes with X that we know about is an arbitrary function of X; *i.e.*,

$$P = -i\,\hbar\,D + f(X) \quad \Longleftrightarrow \quad \langle x\,|P|x'\,\rangle = -i\,\hbar\,\frac{d}{dx}\,\delta(x - x') + f(x)\,\delta(x - x')$$

Rewriting Postulate 2 (cont.)

It turns out that such a redefinition results in no changes in the observable quantities and that one can absorb the change by redefinition of the $\{|x\rangle\}$ basis. Specifically, if we define

$$g(x) = \int_{x_0}^{x} f(x')\,dx'$$

then the redefinition

$$|x\rangle \quad \longrightarrow \quad |\widetilde{x}\rangle = e^{-\frac{i}{\hbar}\,g(x)}|x\rangle$$

results in

$$\langle\widetilde{x}\,|X|\widetilde{x}'\,\rangle = x\,\delta(x - x')$$

Rewriting Postulate 2 (cont.)

Let's see how the redefinition affects the matrix elements of P:

$$
\begin{aligned}
&\langle \widetilde{x} \,|P|\widetilde{x}\,' \rangle \\
&= \langle x \,| e^{\frac{i}{\hbar}\, g(x)} P e^{-\frac{i}{\hbar}\, g(x)} |x' \rangle \\
&= e^{\frac{i}{\hbar}\, g(x)} \int_{-\infty}^{\infty} dx'' \langle x \,|P|x'' \rangle e^{-\frac{i}{\hbar}\, g(x)} \langle x'' \,|x' \rangle \\
&= e^{\frac{i}{\hbar}\, g(x)} \int_{-\infty}^{\infty} dx'' \left[-i\,\hbar \frac{d}{dx} \delta(x - x'') + f(x)\delta(x - x'') \right] e^{-\frac{i}{\hbar}\, g(x)} \delta(x - x'') \\
&= e^{\frac{i}{\hbar}\, g(x)} \left\{ -i\,\hbar \frac{d}{dx} \left[e^{-\frac{i}{\hbar}\, g(x)} \delta(x - x'') \right] + f(x)\delta(x - x') \right\} \\
&= e^{\frac{i}{\hbar}\, g(x)} \left\{ - \left[\frac{d}{dx} g(x) \right] \left[e^{-\frac{i}{\hbar}\, g(x)} \delta(x - x'') \right] - i\,\hbar e^{-\frac{i}{\hbar}\, g(x)} \frac{d}{dx}\, \delta(x - x'') + f(x)\delta(x - x') \right\} \\
&= -i\,\hbar \frac{d}{dx}\, \delta(x - x')
\end{aligned}
$$

That is, by redefining the basis, we recover the original matrix elements for X and P. This is a very interesting freedom that we have not really studied before. This freedom is explored in Shankar Problem 7.4.9. The above kind of transformation is called a *gauge transformation*; you may have seen it in electromagnetism before. The fact that the physics is unchanged under the above gauge transformation can actually be interpreted as rendering the electromagnetic field necessary.

Rewriting Postulate 2 (cont.)

The idea of using canonical commutators to rewrite Postulate 2 is key to connecting it in a unified manner to classical mechanics and to generalizing the postulates to new situations involving new observables. For example, we can use it to rewrite Postulate 2 for situations involving multiple degrees of freedom (*e.g.*, x, y, and z) or multiple particles:

The independent variables $\{x_i\}$ and $\{p_i\}$ that describe completely n degrees of freedom in classical mechanics are represented by Hermitian operators $\{X_i\}$ and $\{P_i\}$ in the Hilbert space of states. The $\{X_i\}$ and $\{P_i\}$ are defined by the canonical commutators

$$[X_i, P_j] = i\,\hbar\,\delta_{ij} \qquad [X_i, X_j] = 0 \qquad [P_i, P_j] = 0$$

Any arbitrary classical dynamical variable $\omega(\{x_i\}, \{p_i\})$ has a corresponding Hermitian operator

$$\Omega(\{X_i\}, \{P_i\}) = \omega(\{x_i \rightarrow X_i\}, \{p_i \rightarrow P_i\})$$

where we simply replace $\{x_i\}$ and $\{p_i\}$ in ω with $\{X_i\}$ and $\{P_i\}$ to obtain $\Omega(\{X_i\}, \{P_i\})$.

Rewriting Postulate 2 (cont.)

To be utterly clear, the original writing of the N-dimensional version of Postulate 2 would tell us that the first commutator is

$$\langle x_1, x_2, \ldots, x_N \,|[X_i, P_j]|x_1', x_2', \ldots, x_N' \,\rangle \\ = i\,\hbar\,\delta_{ij}\,\delta(x_1 - x_1')\,\delta(x_2 - x_2')\cdots\delta(x_N - x_N')$$

This is consistent with our new writing as follows:

$$\langle x_1, x_2, \ldots, x_N \,|[X_i, P_j]|x_1', x_2', \ldots, x_N' \,\rangle = \langle x_1, x_2, \ldots, x_N \,|i\,\hbar\,\delta_{ij}|x_1', x_2', \ldots, x_N' \,\rangle \\ = i\,\hbar\,\delta_{ij}\,\delta(x_1 - x_1')\,\delta(x_2 - x_2')\cdots\delta(x_N - x_N')$$

(The other commutators always vanish, so we do not bother to write them out to the above level of detail.)

To study systems with multipled degrees of freedom or multiple partiles, we also need to deal with Postulate 1, which constructs the Hilbert space. We will do that later. But, clearly, we already have an extension of Postulate 2 to such systems.

Section 7
The Heisenberg Uncertainty Relation

Lecture 24:
The Heisenberg Uncertainty Relation

Date Given: 2008/11/24
Date Revised: 2008/11/24

Review of Expectation Values and Uncertainty

Recall in Section 4.4, in connection with Postulate 3, we defined the expectation value and uncertainty for a physical observable Ω because these are the most definite quantities we can calculate given the probabilistic nature of measurement outcomes in quantum mechanics. They are

$$\langle \Omega \rangle = \langle \psi \, | \Omega | \psi \, \rangle = \sum_i P(\omega_i)\, \omega_i \quad \text{or} \quad \int_{\omega_-}^{\omega_+} d\omega \, P(\omega)\, \omega$$

$$\langle (\Delta\Omega)^2 \rangle = \langle \psi \, | \left[\Omega - \langle \Omega \rangle \right]^2 | \psi \, \rangle = \langle \psi \, | \left[\Omega^2 - \langle \Omega \rangle^2 \right] | \psi \, \rangle = \langle \psi \, | \Omega^2 | \psi \, \rangle - \langle \Omega \rangle^2$$

$$= \sum_i P(\omega_i)\, (\omega_i - \langle \Omega \rangle)^2 \quad \text{or} \quad \int_{\omega_-}^{\omega_+} d\omega \, P(\omega)\, (\omega - \langle \Omega \rangle)^2$$

Derivation of the Uncertainty Relations

When we consider the uncerainties in two variables, we shall see that the product of their uncertainties has a lower limit that is related to their commutator. This should not surprise as, as we already know that commuting operators are simultaneously diagonalizable and hence can simultaneously have vanishing uncertainties. But now we will generically prove the converse case, in which we consider possibly noncommuting operators.

Consider the commutator of the operators corresponding to two physical variables and write it in the following form:

$$[\Omega, \Lambda] = i\,\Gamma$$

If the commutator vanishes, then Γ is the zero operator, the operator that sends every state to the null vector. Because Ω and Λ are Hermitian by assumption, Γ is also Hermitian (which is why the i was introduced).

Deriving the Uncertainty Relation (cont.)

Now, consider the product of the squares of the uncertainties of the two operators for an arbitrary state $|\psi\rangle$:

$$
\begin{aligned}
\langle(\Delta\Omega)^2\rangle\langle(\Delta\Lambda)^2\rangle &= \langle\psi|\widetilde{\Omega}^2|\psi\rangle\langle\psi|\widetilde{\Lambda}^2|\psi\rangle \quad \text{with} \quad \widetilde{\Omega} = \Omega - \langle\Omega\rangle,\ \widetilde{\Lambda} = \Lambda - \langle\Lambda\rangle \\
&= \left(\widetilde{\Omega}^\dagger|\psi\rangle\right)^\dagger \widetilde{\Omega}|\psi\rangle \left(\widetilde{\Lambda}^\dagger|\psi\rangle\right)^\dagger \widetilde{\Lambda}|\psi\rangle \\
&= \left(\widetilde{\Omega}|\psi\rangle\right)^\dagger \widetilde{\Omega}|\psi\rangle \left(\widetilde{\Lambda}|\psi\rangle\right)^\dagger \widetilde{\Lambda}|\psi\rangle \quad \text{bec. } \widetilde{\Omega},\ \widetilde{\Lambda} \text{ are Hermitian} \\
&= \left|\widetilde{\Omega}|\psi\rangle\right|^2 \left|\widetilde{\Lambda}|\psi\rangle\right|^2 \\
&\geq \left|\left(\widetilde{\Omega}|\psi\rangle\right)^\dagger \widetilde{\Lambda}|\psi\rangle\right|^2 \quad \text{by Schwarz Inequality, Equation 3.20} \\
&= \left|\langle\psi|\widetilde{\Omega}\,\widetilde{\Lambda}|\psi\rangle\right|^2 \\
&= \left|\langle\psi|\left\{\frac{1}{2}\left[\widetilde{\Omega}\,\widetilde{\Lambda} + \widetilde{\Lambda}\,\widetilde{\Omega}\right] + \frac{1}{2}\left[\widetilde{\Omega}\,\widetilde{\Lambda} - \widetilde{\Lambda}\,\widetilde{\Omega}\right]\right\}|\psi\rangle\right|^2 \\
&= \left[\frac{1}{2}\langle\psi|\left[\widetilde{\Omega},\widetilde{\Lambda}\right]_+|\psi\rangle + \frac{1}{2}\langle\psi|\left[\widetilde{\Omega},\widetilde{\Lambda}\right]|\psi\rangle\right|^2
\end{aligned}
$$

Deriving the Uncertainty Relation (cont.)

To evaluate the above, we use the facts that the operator $\left[\widetilde{\Omega},\widetilde{\Lambda}\right]_+$ is Hermitian and $\left[\widetilde{\Omega},\widetilde{\Lambda}\right] = [\Omega,\Lambda] = i\,\Gamma$ is anti-Hermitian, and so the expectation values $\langle\psi\,|\left[\widetilde{\Omega},\widetilde{\Lambda}\right]_+|\,\psi\,\rangle$ and $\langle\psi\,|\Gamma|\psi\,\rangle$ are perfectly real. The expression above is then of the form $|a+i\,b|^2$ where a and b are real, so we know the result is a^2+b^2. We thus have

$$\langle(\Delta\Omega)^2\rangle\langle(\Delta\Lambda)^2\rangle \geq \frac{1}{4}\left[\langle\psi\,|\left[\widetilde{\Omega},\widetilde{\Lambda}\right]_+|\,\psi\,\rangle\right]^2 + \frac{1}{4}\left[\langle\psi\,|\Gamma|\psi\,\rangle\right]^2$$

This is the generic uncertainty relation. It is not that useful yet because the right side depends on the state $|\,\psi\,\rangle$.

When the commutator is the canonical value $i\,\Gamma = i\,\hbar$, then the above simplifies to

$$\langle(\Delta\Omega)^2\rangle\langle(\Delta\Lambda)^2\rangle \geq \frac{1}{4}\left[\langle\psi\,|\left[\widetilde{\Omega},\widetilde{\Lambda}\right]_+|\,\psi\,\rangle\right]^2 + \frac{\hbar^2}{4}$$

$$\text{or} \qquad \sqrt{\langle(\Delta\Omega)^2\rangle}\sqrt{\langle(\Delta\Lambda)^2\rangle} \geq \frac{\hbar}{2}$$

where we made the last step because the first term is always nonnegative. This is the Heisenberg uncertainty relation.

Saturation of the Heisenberg Uncertainty Relation

The first condition is

$$\widetilde{\Omega}|\psi\rangle = c\,\widetilde{\Lambda}\,|\psi\rangle$$

in order that the Schwarz inequality used early in the proof be saturated. Note that we are not requiring the relation $\widetilde{\Omega} = c\,\widetilde{\Lambda}$ to hold in general – then the two would commute and this would all be trivial. We are simply requiring that this hold for the particular state $|\psi\rangle$ that is going to be a state that saturates the inequality.

The second condition is for the first term in the generic uncertainty relation to vanish:

$$\langle\psi|\left[\widetilde{\Omega},\widetilde{\Lambda}\right]_+|\psi\rangle = 0$$

This is obvious, as if this term is nonzero, then it ensures that the relation cannot be an equality.

Example 7.1: The Gaussian Wavefunction

We have twice gone through the demonstration that a state with Gaussian $\{|x\rangle\}$-basis representation always saturates the Heisenberg uncertainty relation for X and P, giving

$$\sqrt{\langle(\Delta X)^2\rangle}\sqrt{\langle(\Delta P)^2\rangle} = \frac{\hbar}{2}$$

We studied this for both a wave packet propagating freely and for the simple harmonic oscillator. In Section 9.3, Shankar shows explicitly that, for any potential, the Gaussian wavefunction is *the only* state that renders the inequality an equality by using the first condition above to obtain a differential equation that determines the wavefunction,

$$(P - \langle P\rangle)\,|\psi\rangle = c\,(X - \langle X\rangle)\,|\psi\rangle$$
$$\left(-i\,\hbar\,\frac{d}{dx} - \langle P\rangle\right)\psi(x) = c\,(x - \langle X\rangle)\,\psi(x)$$

where, in going from the first line to the second, we took the product with $\langle x|$ on the left, inserted completeness in the $\{|x\rangle\}$, and did the completeness integral. The second condition from above is used in Shankar's proof, too. It is worth going through Shankar's proof for the sake of the technique.

Examples (cont.)

Example 7.2: Hydrogen Atom

Shankar goes through in detail a calculation of the ground state energy and radius of the hydrogen atom. Again, it is worth studying the technique used, in particular the way in which he approximates the potential term in the Hamiltonian, which is not trivially written as a function of $\langle(\Delta X)^2\rangle$, and then differentiates E with respect to $\langle(\Delta X)^2\rangle$ to find the minimum possible energy.

Example 7.3: Diffraction at a Screen

Consider a particle traveling in the x direction with momentum $\hbar k$ incident on a screen with an aperture extending from $y = -a$ to $y = a$. The particle's position-space wavefunction to the left of the screen is e^{ikx}; there is no y dependence. The aperture truncates the wavefunction in y so it vanishes outside the interval $[-a, a]$. The y position uncertainty then becomes $\sqrt{\langle (\Delta Y)^2 \rangle} = \frac{a}{\sqrt{3}}$ (you can check this calculation easily). So the y momentum uncertainty becomes

$$\sqrt{\langle (\Delta P_y)^2 \rangle} \geq \frac{\hbar}{2} \frac{1}{\sqrt{\langle (\Delta Y)^2 \rangle}} = \frac{\hbar\sqrt{3}}{2a}$$

Thus, the propagating plane wave, which initially had no y momentum, acquires a rms y momentum of this size. This causes the wavefunction to spread out in y; the angular extent that the image of the particle beam on a far screen will cover is approximately

$$\sqrt{\langle (\Delta\theta)^2 \rangle} = \frac{\sqrt{\langle (\Delta P_y)^2 \rangle}}{\langle P_x \rangle} \geq \frac{\frac{\hbar\sqrt{3}}{2a}}{\hbar k} = \frac{\sqrt{3}}{2ka}$$

Example 7.4: Size of Nuclei

It is experimentally observed that the binding energy of nuclei is in the few MeV/nucleon range; meaning that nuclei can be caused to break apart by interactions with photons or other particles having this amount of energy. This information can be used to determine the approximate size of a nucleus via simple particle-in-a-box type considerations.

Let α be the typical binding energy per nucleon. Then α is a lower limit on the depth of the potential well, and thus an upper limit on the energy of each nucleon. We may get the momentum uncertainty of a single nucleon from the energy via

$$\frac{\langle (\Delta P)^2 \rangle}{2\, m_p} = \alpha \quad \Longrightarrow \quad \langle (\Delta P)^2 \rangle = 2\, \alpha\, m_p$$

From this, let's use the uncertainty principle to determine the position uncertainty

$$\langle (\Delta X)^2 \rangle = \frac{\hbar^2}{4} \frac{1}{\langle (\Delta P)^2 \rangle} = \frac{\hbar^2}{4} \frac{1}{2\, \alpha\ m_p}$$

Examples (cont.)

Numerically, we have

$$\begin{aligned}\sqrt{\langle(\Delta X)^2\rangle} &= \frac{\hbar}{2A\sqrt{2\,\alpha\,m_p}} \\ &= \frac{1.0\times10^{-34}\ \text{J s}}{2A\sqrt{2\,\alpha\times10^{6}\times1.6\times10^{-19}\,\text{J}\times1.7\times10^{-27}\,\text{kg}}} \\ &= \frac{2.1\times10^{-15}\,\text{m}}{\sqrt{\alpha}}\end{aligned}$$

where we have converted α to J to do the calculation. For most nuclei, $\alpha \approx 8$ MeV/nucleon, so we get 0.7 fm. Now, this scaling with α should not necessarily be believed – bigger nuclei have higher binding energies but are also bigger – but the order of magnitude is correct. In practice, nuclei have radii that follow $r = 1.2\,A^{1/3}$ fm.

Lecture 25:
The Energy-Time Uncertainty Relation

Date Given: 2008/11/26
Date Revised: 2008/11/26

H and $i\hbar\frac{\partial}{\partial t}$ as conjugate operators

The Schrödinger Equation tells us

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle$$

Let's take the matrix element on the left with $\langle x|$ and insert completeness:

$$i\hbar\frac{\partial}{\partial t}\psi_x(x,t) = \int_{-\infty}^{\infty} dx'\langle x|H|x'\rangle\psi_x(x',t)$$

(The total time derivative now becomes a partial time derivative because of the appearance of x having projected onto the $\{|x\rangle\}$ basis.) Notice the striking similarity to the expression for the P operator acting in the $\{|x\rangle\}$ basis:

$$-i\hbar\frac{\partial}{\partial x}\psi_x(x,t) = \int_{-\infty}^{\infty} dx'\langle x|P|x'\rangle\psi_x(x',t)$$

where again a partial derivative appears. The variables t and E are mathematically conjugate to each other in a manner similar to x and p. However, this is not rigorous in the sense of operators because there is no t operator!

The Energy-Time Uncertainty Relation (cont.)

We can see the conjugate relationship more rigorously by considering a free-particle wave packet with

$$\langle X\rangle = \frac{p_0}{m}\,t \qquad \langle(\Delta X)^2\rangle_{t=0} = \sigma_x^2 \qquad \langle P\rangle = p_0 \qquad \langle(\Delta P)^2\rangle = \sigma_p^2 = \frac{\hbar^2}{4\,\sigma_x^2}$$

with $\sigma_p/p \ll 1$. The wavefunction as a function of position and time is

$$\begin{aligned}
\psi_x(x,t) &= \langle x\,|\psi(t)\,\rangle = \langle x\,|U(t)\,|\psi(t=0)\,\rangle \\
&= \int_0^\infty dp\, e^{-\frac{i}{\hbar}\,\frac{p^2}{2\,m}\,t}\,[\langle x\,|E_p\,\rangle\langle E_p\,|\psi(t=0)\,\rangle + \langle x\,|E_{-p}\,\rangle\langle E_{-p}\,|\psi(t=0)\,\rangle] \\
&\approx \frac{1}{\sqrt{2\,\pi\,\hbar}}\int_0^\infty dp\, e^{-\frac{i}{\hbar}\,\frac{p^2}{2\,m}\,t}\,e^{\frac{i}{\hbar}\,p\,x}\psi_p(p)
\end{aligned}$$

where we have run the integral from $-\infty$ to ∞ because $\psi_p(p)$ is strongly peaked near p_0 and $p_0/\sigma_p \gg 1$ so that the integrand is very small near the origin and for $p < 0$.

The Energy-Time Uncertainty Relation (cont.)

Now, let's specialize to $x = 0$: we want to see what the time dependence of the wavefunction at the origin is. We have

$$\begin{aligned}\psi_x(0,t) &\approx \frac{1}{\sqrt{2\,\pi\,\hbar}} \int_0^\infty dp\, e^{-\frac{i}{\hbar}\,\frac{p^2}{2\,m}\,t} \psi_p(p) \\ &\approx \frac{1}{\sqrt{2\,\pi\,\hbar}} \int_{-\infty}^\infty dE\, e^{-\frac{i}{\hbar}\,E\,t}\, \psi_E(E) \\ &\text{where} \quad \psi_E(E) = \sqrt{\frac{m}{2E}}\, \psi_p(p) = \frac{m}{p}\, \psi_p(p)\end{aligned}$$

This looks like a Fourier transform in E! We are thus led to the idea that E and t should be conjugate variables in the same way that x and p are, and thus that there might be an uncertainty relation between them.

The Energy-Time Uncertainty Relation (cont.)

Let's now compare the width in E and the width in t at the position x. Recall that we calculated the width in E for a propagating free-particle Gaussian wave packet in Section 5.5 and found it to be

$$\sqrt{\langle (\Delta E)^2 \rangle} \approx 2\,E\,\frac{\sigma_p}{p_0} = \frac{p_0^2}{m}\,\frac{\sigma_p}{p_0} = \frac{p_0}{m}\,\sigma_p$$

At a given point x, what is the width in t? It is related to the width in x by the expectation value of the velocity:

$$\sqrt{(\Delta t)^2} = \frac{\sqrt{\langle (\Delta X)^2 \rangle}}{v} = \frac{\sigma_x(t)}{p_0/m} \geq \frac{\sigma_x}{p_0/m}$$

where we remind the reader that $\sigma_x(t) \geq \sigma_x$ because the width in position space of the free-particle wave packet grows with time. This quantity is essentially the uncertainty in the time at which the wave packet passes through the point x. Note that we have not written the width in t as an expectation value because there is no t operator.

The Energy-Time Uncertainty Relation (cont.)

Multiplying the two uncertainties together, we have

$$\sqrt{\langle(\Delta E)^2}\sqrt{(\Delta t)^2} \geq \frac{p_0}{m}\,\sigma_p\,\frac{\sigma_x}{p_0/m} = \sigma_p\,\sigma_x = \frac{\hbar}{2}$$

Hence, we obtain a time-energy uncertainty relation. Note the very special way we had to derive this expression – it was hardly generic. But something like this usually holds in general. For eigenstates, the energy uncertainty vanishes, but the time uncertainty becomes infinite because the particle must be in that eigenstate for all time in order for its energy to be truly perfectly defined.

The Energy-Time Uncertainty Relation (cont.)

Other Examples of the Energy-Time Uncertainty Relation

- **Decaying states**
  We have not yet discussed Hamiltonians that can result in the decay of states – *e.g.*, decay of an excited atom to the ground state, decay of a particle. But, a similar uncertainty relation holds for them. The classic example is the energy width of a short-lived fundamental particle such as the Z boson. Though the Z has a well-defined mass $M_Z = 91.2$ GeV/c^2, it has a mass width of about $\Gamma_z = 2.5$ GeV/c^2. That is, two electrons collided together with total center of mass energy a bit less or a bit more than M_Z can still create a Z at rest. This is because the Z decays with a lifetime of $\tau = \hbar/2\,\Gamma_z$, which, when put into the right units, is about 10^{-25} s. One can prove this rigorously by decomposing the wavefunction of the Z into mass eigenstates and seeing that the wavefunction has a spread in energy of Γ_z.

The Energy-Time Uncertainty Relation (cont.)

- **Virtual Particles**
  In particle physics, energy conservation can be violated on very short timescales by "virtual processes". For example, the scattering of two particles via transfer of an intermediate particle (a photon, a Z, etc.) can be though of as emission of the mediating particle followed by absorption a time Δt later. For the short life of the mediating particle, it need not conserve energy – it can have an energy that is not tied to its momentum by its rest mass. Momentum must be conserved at every interaction point, but energy only need be conserved over the entire process. The extent to which energy can be lost or gained temporarily is related to the time between scatters. We will see a simple version of this when we consider second-order perturbation theory later.

Section 8
Semiclassical Limit

Lecture 26:
The Semiclassical (WKB) Approximation

Date Given: 2008/11/26, 2008/12/01, 2008/12/03
Date Revised: 2008/12/03

Derivation for Unbound States

Motivation

Consider a particle in one dimension in a region of constant potential energy V extending from 0 to L. (The generalization to an arbitrary position away from the origin is straightforward and introduces only a constant multiplicative factor.) We know that, in this region, the solution to the Schrödinger Equation takes the form

$$\psi(x) = \psi(x=0)\, e^{\pm \frac{i}{\hbar} p x} \qquad p = \sqrt{2\, m\, (E - V)}$$

Remember that the form of the potential outside the interval $[0, L]$ only affects the boundary conditions at 0 and L, which determine the coefficients of the $+$ and $-$ solutions; the functional form of the solution is independent of the potential elsewhere.

Derivation for Unbound States (cont.)

Now, let's try to generalize the plane-wave solution to a potential that depends continuously on position. Any reasonable such potential may be treated as the limit of a set of piecewise constant potentials, and our plane-wave solution is strictly valid in each region. But the parameters p and λ will change from region to region. That is, we may write the solution in the following piecewise form, where x_1, x_2, etc., denote the boundaries and p_1, p_2, etc. denote the value of p in the region beginning at x_1, x_2, etc.:

$$\psi(x) = \begin{cases} \psi(x=0)\, e^{\pm \frac{i}{\hbar}\, p_0\, x} & 0 \le x < x_1 \\ \psi(x=x_1)\, e^{\pm \frac{i}{\hbar}\, p_1\, (x-x_1)} & x_1 \le x < x_2 \\ \psi(x=x_2)\, e^{\pm \frac{i}{\hbar}\, p_2\, (x-x_2)} & x_2 \le x < x_3 \\ \vdots & \vdots \end{cases}$$

We may write the above solution in the following less cumbersome form

$$\psi(x) = \psi(x=0)\, \exp\left(\pm \frac{i}{\hbar} \int_0^x dx'\, p(x')\right) \quad \text{where} \quad p(x) = \sqrt{2\, m\, (E - V(x))} \tag{8.1}$$

where $p(x)$ is now a piecewise constant function in the same way that $V(x)$ is; $p = p_0$ for $0 \le x < x_1$, $p = p_1$ for $x_1 \le x < x_2$, etc.

Derivation for Unbound States (cont.)

Now, let's take the limit of $V(x)$ becoming continuous. The definition of $p(x)$ is still valid, but the solution is no longer rigorously correct (try plugging in to the Schrödinger Equation and you will see this). But we expect the solution to be approximately correct if $V(x)$ varies slowly enough. What is meant by slowly enough? The integral in the above form is just the integrated phase accumulated from 0 to x. The phase shift per unit length is given by the wavelength λ where $\lambda = 2\pi\hbar/p$. When the potential is piecewise constant, then λ is piecewise constant. We thus expect that the above solution is valid if the fractional variation in λ over one λ is small compared to 1:

$$\left|\frac{1}{\lambda}\left(\frac{d\lambda}{dx}\lambda\right)\right| \ll 1 \quad \Longrightarrow \quad \left|\frac{d\lambda}{dx}\right| \ll 1$$

Derivation for Unbound States (cont.)

Derivation (for Unbound States)

Let's prove the above more rigorously. Our expected solution is of the form

$$\psi(x) = A\, e^{\frac{i}{\hbar}\phi(x)}$$

where $\phi(x)$ may be complex to avoid restricting the form of the solution. The Schrödinger Equation for our problem is

$$-\frac{\hbar^2}{2\,m}\frac{d^2}{dx^2}\,\psi(x) + V(x)\,\psi(x) = E\,\psi(x)$$

which we may rewrite as

$$\left[\frac{d^2}{dx^2} + \frac{p^2(x)}{\hbar^2}\right]\psi(x) = 0$$

Now, inserting our expected solution, we have

$$\left(\frac{1}{\hbar}\,\frac{d\phi}{dx}\right)^2 - \frac{i}{\hbar}\,\frac{d^2\phi}{dx^2} - \frac{p^2}{\hbar^2} = 0$$

Derivation for Unbound States (cont.)

Now, we would like to apply our approximation. We want to apply our criterion $|d\lambda/dx| \ll 1$ directly. The instantaneous wavelength of the solution is the distance over which the phase changes by 2π, which we can calculate from $\phi(x)$ via

$$2\pi = \frac{1}{\hbar}\left[\phi(x+\lambda) - \phi(x)\right]$$

$$2\pi\hbar \approx \lambda \frac{d\phi}{dx}$$

$$\lambda(x) \approx \frac{2\pi\hbar}{\frac{d\phi}{dx}}$$

Now, let's calculate the derivative:

$$\frac{d\lambda}{dx} \approx -\frac{2\pi\hbar}{\left(\frac{d\phi}{dx}\right)^2}\frac{d^2\phi}{dx^2}$$

Derivation for Unbound States (cont.)

So our condition $|d\lambda/dx| \ll 1$ is

$$\left| \frac{2\,\pi\,\hbar}{\left(\frac{d\phi}{dx}\right)^2} \frac{d^2\phi}{dx^2} \right| \ll 1 \quad \Longleftrightarrow \quad \left| \frac{1}{\hbar} \frac{d^2\phi}{dx^2} \right| \ll \frac{1}{2\,\pi} \frac{1}{\hbar^2} \left(\frac{d\phi}{dx}\right)^2$$

This tells us that the second term in the differential equation is small compared to the first.

For convenience, let's define α to be the ratio of the two terms, which is just $d\lambda/dx$:

$$\alpha \equiv \frac{\frac{1}{\hbar} \frac{d^2\phi}{dx^2}}{\frac{1}{2\,\pi} \frac{1}{\hbar^2} \left(\frac{d\phi}{dx}\right)^2} = \frac{d\lambda}{dx}$$

Our approximation condition is therefore $|\alpha| \ll 1$.

Derivation for Unbound States (cont.)

However, we may not just ignore the second term because the result will then be too gross an approximation. For the sake of pedagogy, let's go through it to see that. The differential equation for ϕ becomes

$$\left(\frac{1}{\hbar}\frac{d\phi}{dx}\right)^2 - \frac{p^2}{\hbar^2} = 0 \quad \Longleftrightarrow \quad \frac{d\phi}{dx} = \pm p(x)$$

Integrating, we obtain

$$\phi(x) = \int^x dx' p(x')$$

Our solution implies $d^2\phi/dx^2 = \pm dp/dx$. But our approximation has consisted of ignoring $d^2\phi/dx^2$. This implies we should ignore dp/dx, which is equivalent to ignoring $dV/dx = 0$. But p will then just be a constant. This is too gross an approximation, we have completely ignored the spatial dependence of the potential.

Derivation for Unbound States (cont.)

So, rather than ignoring terms of order α, let us carefully expand ϕ to first order in α; our approximation condition requires that $|\alpha| \ll 1$, so it is a good parameter to expand in terms of. That is, let

$$\phi(x) = \phi_0(x) + \alpha\,\phi_1(x)$$

and, similarly,

$$\frac{d}{dx}\,\phi(x) = \frac{d}{dx}\,\phi_0(x) + \alpha\,\frac{d}{dx}\,\phi_1(x)$$
$$\frac{d^2}{dx^2}\,\phi(x) = \frac{d^2}{dx^2}\,\phi_0(x) + \alpha\,\frac{d^2}{dx^2}\,\phi_1(x)$$

Note that this is not a Taylor expansion in position; α is just a dimensionless parameter. You can think of it as a Taylor expansion in terms of α where higher order terms in α give us a better approximation to the exact wavefunction.

Derivation for Unbound States (cont.)

The condition $|\alpha| \ll 1$ actually puts a condition on the expansion coefficients, which we can see by simply writing the definition of α in terms of the expansions:

$$\alpha \equiv \frac{\frac{1}{\hbar}\,\frac{d^2\phi}{dx^2}}{\frac{1}{2\,\pi}\,\frac{1}{\hbar^2}\left(\frac{d\phi}{dx}\right)^2} = \frac{\frac{1}{\hbar}\left(\frac{d^2\phi_0}{dx^2} + \alpha\,\frac{d^2\phi_1}{dx^2}\right)}{\frac{1}{2\,\pi}\,\frac{1}{\hbar^2}\left(\frac{d\phi_0}{dx} + \alpha\,\frac{d\phi_1}{dx}\right)^2} \approx \frac{\frac{1}{\hbar}\,\frac{d^2\phi_0}{dx^2}}{\frac{1}{2\,\pi}\,\frac{1}{\hbar^2}\left(\frac{d\phi_0}{dx}\right)^2}$$

where, since $|\alpha| \ll 1$ is already a requirement, we have dropped all terms of order α or higher on the right side. This relationship between $\frac{d^2\phi_0}{dx^2}$ and $\left(\frac{d\phi_0}{dx}\right)^2$ will be important below.

Plug the α expansion formulae into our differential equation:

$$\left(\frac{\frac{d\phi_0}{dx} + \alpha\,\frac{d\phi_1}{dx}}{\hbar}\right)^2 - i\,\frac{\frac{d^2\phi_0}{dx^2} + \alpha\,\frac{d^2\phi_1}{dx^2}}{\hbar} - \frac{p^2(x)}{\hbar^2} = 0$$

$$\left(\frac{1}{\hbar}\,\frac{d\phi_0}{dx}\right)^2 + \frac{2\,\alpha}{\hbar^2}\,\frac{d\phi_0}{dx}\,\frac{d\phi_1}{dx} - \frac{i}{\hbar}\,\frac{d^2\phi_0}{dx^2} - \frac{p^2(x)}{\hbar^2} = 0$$

where we have dropped all terms of order α^2. Based on our rewriting of α above, we saw that the third term is already of order α without having to keep the $\frac{d^2\phi_1}{dx^2}$ term; hence the importance of that rewriting.

Derivation for Unbound States (cont.)

The above differential equation consists of terms of zeroth order in α and of first order; that is, we have a polynomial in α. For the equation to be satisfied for arbitrary α, the coefficient of each power of α must vanish. We thus obtain two equations:

$$\left(\frac{1}{\hbar}\frac{d\phi_0}{dx}\right)^2 - \frac{p^2}{\hbar^2} = 0 \qquad \frac{2\,\alpha}{\hbar^2}\frac{d\phi_0}{dx}\frac{d\phi_1}{dx} = \frac{i}{\hbar}\frac{d^2\phi_0}{dx^2}$$

The first equation is the same one we found when we made the approximation $\alpha = 0$; it yields

$$\frac{d\phi_0}{dx} = \pm p(x) \quad \Longleftrightarrow \quad \phi_0 = \pm \int^x dx' p(x')$$

We see how our previous solution was indeed too gross an approximation: it consisted of setting $\alpha = 0$, neglecting any variation in the potential.

Derivation for Unbound States (cont.)

The second equation is

$$\frac{2\,\alpha}{\hbar^2}\,\frac{d\phi_0}{dx}\,\frac{d\phi_1}{dx} = \frac{i}{\hbar}\,\frac{d^2\phi_0}{dx^2}$$

$$-\frac{2\,i\,\alpha}{\hbar}\,\frac{d\phi_1}{dx} = \frac{\frac{d^2\phi_0}{dx^2}}{\frac{d\phi_0}{dx}}$$

Now, here we use our solution for ϕ_0, $\frac{d\phi_0}{dx} = \pm p(x)$ (the $\pm$ sign cancels out in the numerator and denominator of the right side):

$$-\frac{2\,i\,\alpha}{\hbar}\,\frac{d\phi_1}{dx} = \frac{\frac{dp}{dx}}{p}$$

$$-\frac{2\,i\,\alpha}{\hbar}\,\phi_1 + c = \ln p$$

$$\frac{i}{\hbar}\,\alpha\,\phi_1 = -\frac{1}{2}\,\ln\frac{d\phi_0}{dx} + C = \ln\frac{1}{\sqrt{p}} + C$$

Note that we may assume α is constant for the purpose of doing the integral because, by assumption, any variations in α will be of order α^2 and may be ignored.

Derivation for Unbound States (cont.)

So, we have, to first order in α

$$\phi(x) \approx \phi_0(x) + \alpha\,\phi_1(x) = \pm \int^x dx'\,p(x') + \frac{\hbar}{i} \ln \frac{1}{\sqrt{p(x)}} + C$$

$$\psi(x) = A\,e^{\frac{i}{\hbar}\,\phi(x)} = \frac{A}{\sqrt{p(x)}} \exp\left(\pm \frac{i}{\hbar} \int^x dx'\,p(x')\right)$$

The lower limit of integration is arbitrary and its value just modifies the overall phase of the wavefunction. So, we see that the solution follows roughly the form that we had guessed initially, but there is an additional $\frac{1}{\sqrt{p}}$ factor out front. The heuristic explanation of this factor is that it yields $P(x) = |\psi(x)|^2 = \frac{|A|^2}{p(x)} = \frac{|A|^2}{m\,v(x)}$. The mass factor is just an overall constant that is unimportant, but the $v(x)$ factor is position-dependent and simply recalls the classical fact that the dwell time in a small interval about the point x is inversely proportional to the velocity at that point $v(x)$. This is similar to the point we made about the higher excited levels in the SHO (see Section 6.2), where the probability density also becomes inversely proportional to the particle velocity.

The above result for the wavefunction is called the semiclassical approximation because of the assumption that the fractional wavelength change per wavelength is slow. It is also called the Wentzel-Kramers-Brillouin or WKB approximation after its originators.

Derivation for Unbound States (cont.)

One thing we need to check is the self-consistency of the approximation; that is, does it turn out that $|d\lambda/dx| \ll 1$ as we required to start with? We just calculate it directly from the solution. (Note that the $p^{-1/2}$ factor has no effect on the wavelength of the solution, just on the normalization.) But this is trivial to check, by construction. From the final solution, it holds that

$$\lambda(x) \approx \frac{2\,\pi\,\hbar}{\frac{d\phi}{dx}}$$

$$\frac{d\lambda}{dx} \approx \frac{\frac{1}{\hbar}\,\frac{d^2\phi}{dx^2}}{\frac{1}{2\,\pi\,\hbar}\left(\frac{d\phi}{dx}\right)^2} = \alpha$$

So the approximation is self-consistent – if we assume $|d\lambda/dx| \ll 1$ to start with and then calculate $d\lambda/dx$ from the solution, we recover $|d\lambda/dx| \ll 1$. This may seem tautological, but that's the beauty of doing the expansion in terms of α: the self-consistency is manifest (obvious) at all points during the calculation, so in the end it is trivial that self-consistency is assured.

Derivation for Unbound States (cont.)

Regime of Validity in Terms of Classical Quantities

Moreover, the approximation is not too gross anymore, as we now allow $d\lambda/dx \neq 0$, which implies $dV/dx \neq 0$. In fact, let's rewrite our condition on λ in terms of V and E. We have

$$V(x) = E - \frac{p^2(x)}{2\,m} = E - \frac{1}{2\,m}\left(\frac{d\phi_0}{dx}\right)^2 \qquad \frac{dV}{dx} = -\frac{1}{m}\,\frac{d\phi_0}{dx}\,\frac{d^2\phi_0}{dx^2}$$

$$\alpha = \frac{\frac{1}{\hbar}\,\frac{d^2\phi_0}{dx^2}}{\frac{1}{2\,\pi}\,\frac{1}{\hbar^2}\left(\frac{d\phi_0}{dx}\right)^2} = -\frac{\frac{m}{\hbar}\,\frac{dV}{dx}}{\frac{1}{2\,\pi}\,\frac{1}{\hbar^2}\left(\frac{d\phi_0}{dx}\right)^3} = -\frac{\frac{dV}{dx}}{\frac{p^3(x)}{2\,\pi\,\hbar\,m}} = -\frac{\frac{dV}{dx}\,\frac{2\,\pi\,\hbar}{p}}{2\,T}$$

The final expression is essentially the ratio of the work done on the particle by the potential over a distance λ compared to its kinetic energy; the former is the negative of the change in kinetic energy over the same distance. So the semiclassical criterion is simply that the fractional change in kinetic energy over a wavelength be small compared to unity. $\hbar$ comes into the problem by giving the conversion between the classical momentum and the quantum wavelength. Note that it is not simply a statement about the potential itself, which is why the semiclassical approximation is not expected to work for any arbitrary state in a potential; rather, we expect it to work for excited states in which T is large compared to the work that can be done per unit wavelength. dV/dx does not change with the kinetic energy, but T gets bigger and λ gets smaller as T increases.

Derivation for Unbound States (cont.)

Now that we have the semiclassical requirement in an easier-to-understand form, let's discuss under what conditions it is satisfied. First, the condition makes no sense in regions where $E < V(x)$ because then T is negative and there is no wavelength to speak of. We will need to modify the solution for these regions, though we shall see that a reasonable solution can be found. Second, the solution will fail spectacularly when $E \sim V(x)$, *i.e.*, near classical turning points. At these points, $T \longrightarrow 0$, but the numerator is nonzero (in order for there to be a turning point!), so the approximation simply fails. This is equivalent to $p(x) \longrightarrow 0$, in which case the wavefunction's normalizing factor blows up and the dwell time becomes infinite. Both of these problems affect bound states near and outside the classical turning points. We shall deal with these problems in the next section.

Probability Current

We can calculate the probability current in the WKB approximation and see that it is sensible:

$$\begin{aligned} J &= -\frac{i}{2}\,\frac{\hbar}{m}\left(\psi^*(x)\,\frac{\partial}{\partial x}\,\psi(x) - \psi(x)\,\frac{\partial}{\partial x}\,\psi^*(x)\right) \\ &= -\frac{i}{2}\,\frac{\hbar}{m}\left(\pm 2\,\frac{i}{\hbar}\,p(x)\right)\psi^*(x)\,\psi(x) \\ &= P(x)\,v(x) \end{aligned}$$

where $P(x) = \psi^*(x)\,\psi(x)$ is the probability density and $v(x)$ is the velocity. This makes sense: the probability current is simply the probability times the speed at which the particle moves. (Note that the terms involving derivatives of the $p^{-1/2}$ terms cancel because they are the same for ψ and ψ^*.) The simple form arises because of the way $p(x)$ appears in the argument of the complex exponential.

Derivation for Unbound States (cont.)

Our derivation differs somewhat from Shankar's, and this deserves some explanation. The difference lies in how we expand $\phi(x)$. First, Shankar says the expansion parameter should be $\hbar$, using the heuristic argument that $\lambda \to 0$ as $\hbar \to 0$ and that this is equivalent to the desired limit $|d\lambda/dx| \ll 1$. This is not very rigorous. We have instead explicitly made use of the criterion $|d\lambda/dx| \ll 1$, which we derived in a reasonably rigorous manner. Second, Shankar expands in terms of the dimensionful parameter $\hbar$. This is very bad: the statement "$\hbar$ is small" has no meaning because the units of $\hbar$ are arbitrary. In SI units, yes, $\hbar$ seems small, but in units where $\hbar = 1$, $\hbar$ is not small! We have instead expanded in terms of the dimensionless parameter α. The criterion for α to be small, $|\alpha| \ll 1$, ensures mathematical convergence of the power series expansion in α. And the $|\alpha| \ll 1$ criterion follows directly from the requirement $|d\lambda/dx| \ll 1$. Third, our expansion and derivation are self-consistent by construction, whereas Shankar had to check this explicitly (Equations 16.2.14 and 16.2.15).

The point about doing the expansion of $\phi(x)$ in terms of a dimensionless parameter is particularly important and holds for all physics: whenever you do an expansion to make an approximation, your expansion parameter must be dimensionless; otherwise, you have no idea what is meant by the "smallness" of the parameter and the resulting validity of the expansion. All physical units are to some extent arbitrary; the only truly valid criterion for something to be small is for it to be dimensionless and small compared to unity.

Derivation for Unbound States (cont.)

Relation of the WKB Approximation to Classical Mechanics

The phase factor in the WKB approximation is

$$\phi_r(x) = \int^x dx' p(x')$$

(The $_r$ subscript indicates the real part of ϕ; recall that the imaginary part of ϕ gave the $p^{-1/2}$ normalization factor.) Let's consider the corresponding classical problem. In that limit, p is the classical momentum. Let's assume there is a classical path, $x_{cl}(t)$, that solves the Euler-Lagrange equation for the problem. We may do a change of variables to t, rewriting the above integral as

$$\phi_r(x_{cl}(t)) = \int^t dt \frac{dx_{cl}}{dt} p(x_{cl}(t))$$

The meaning of the notation is: if you want to know $\phi_r(x)$, calculate the t this point x corresponds to along the classical path $x_{cl}(t)$ and evaluate the integral with that t as the upper limit, using $x_{cl}(t)$ in the integrand. We now recall from classical mechanics that, for any path (classical or not), it holds that

$$H = \frac{dx}{dt} p - L$$

Derivation for Unbound States (cont.)

We may then rewrite

$$\phi_r(x_{cl}(t)) = \int^t dt\, L_{cl} + \int^t dt\, H_{cl}$$

where we have put $_{cl}$ subscripts on L and H because we have specified the path $x_{cl}(t)$; L and H are no longer the generic Lagrangian and Hamiltonian functions, which may be evaluated for any path, classical or not. Since H_{cl} is evaluated for the classical path that corresponds to energy E (recall the definition of $p(x)$), we know that H_{cl} takes on the value E at all times and we may replace it, yielding

$$\phi_r(x_{cl}(t)) = \int^t dt\, L_{cl} + E\, t$$

We thus see that $\phi_r(x_{cl}(t))$ is essentially the classical action $S_{cl} = \int dt\, L_{cl}$ evaluated along the classical path with energy E that would arrive at point x_{cl} at time t. The $E\, t$ factor becomes irrelevant when incorporated in the wavefunction, as it is just a time-evolving complex phase factor with no position dependence; it does not affect the probability density or the momentum. It must be in fact canceled out by a corresponding term in the action integral in order for the result to have no explicit time dependence, as is necessary for ϕ_r to be a function of x_{cl} only.

Derivation for Unbound States (cont.)

Thus, the wavefunction is

$$\psi(x) = \frac{A}{\sqrt{p(x)}}\, e^{\frac{i}{\hbar}\phi_r(x)} = \frac{A}{\sqrt{p(x_{cl}(t))}}\, \exp\left(\frac{i}{\hbar}\left[S_{cl}(x) + E\, t\right]\right)$$

where t should be understood as being related to x through the classical path $x_{cl}(t)$ and $S_{cl}(x)$ corresponds to evaluating the classical action along the classical path that passes through x at the time t at which $x_{cl}(t) = x$. The appearance of the classical path $x_{cl}(t)$ and action S_{cl} in such prominent roles reflects the semiclassical nature of the WKB approximation. In deriving the above relation between the WKB approximation and the classical action, we have scratched the surface of the ties between classical and quantum mechanics. If we have time later, we will discuss the path integral formalism, which makes explicit use of the action, using $e^{\frac{i}{\hbar}S}$ as a weighting factor for all paths to define the propagator and the wavefunction. Taking the classical limit of the path integral formalism results in a differential equation for S_{cl} that reproduces the Hamilton-Jacobi equation of classical mechanics. We note that the quantity $\int^x dx' p(x')$ is known as Hamilton's Characteristic Function or the abbreviated action in classical mechanics.

Bound state difficulties and how to deal with them

As we noted above, for bound states, the WKB approximation remains valid in the classically allowed region far from the turning points, but falls apart near and outside the classical turning points. This is not too surprising, as these are fundamentally quantum-mechanical regions.

For the classically unallowed region, $E < V(x)$, our original motivation for the WKB solution remains true with the complex exponential in $p(x)$ replaced by a real exponential in $\kappa(x) = \sqrt{2\,m\,(V(x) - E)}$. Moreover, if we replace $\lambda(x)$ by the wavefunction decay length modulo $2\,\pi$, $\lambda(x) = 2\,\pi\,\hbar/p(x) \rightarrow \gamma(x) = 2\,\pi\,\hbar/\kappa(x)$, assume a wavefunction of the form $e^{\pm\chi(x)}$, and make the requirement $|d\gamma/dx| \ll 1$, we find that things carry through in a straightforward way. The result will be that

$$\psi(x) = \frac{A}{\sqrt{\kappa(x)}}\,\exp\left(\pm\frac{1}{\hbar}\int^{x} dx'\,\kappa(x')\right)$$

where the sign needs to be chosen depending on whether the region extends to $\pm\infty$; both signs are valid if the classically unallowed region is finite.

Derivation for Bound States (cont.)

For the region around the classical turning point, we still have a problem because both $|d\lambda/dx| \ll 1$ and $|d\gamma/dx| \ll 1$ fail to hold: the wavelength and decay constant become infinite as one approaches such a point, so the approximation fails. One simply cannot approximate the solution to the Schrödinger Equation here. One must solve the Schrödinger Equation completely in this region. If the potential varies smoothly enough, one may be able to approximate it by a linear function, or perhaps a polynomial, to make this easier. In the worst case, one can solve the Schrödinger Equation numerically in this small interval. We term this interval the transition region.

Finally, one simply joins the solutions in the various regions. By assumption, the potential is smoothly varying at the boundaries between the transition region and the classically allowed and unallowed regions. Therefore, based on similar analyses we did for one-dimensional problems with finite step potentials, the wavefunction and its derivative must be continuous. We thus obtain matching conditions that set the coefficients in the various regions, yielding a complete solution.

One might ask: what is the benefit of WKB if we still need to solve the Schrödinger Equation explicitly or numerically in the transition region? The reason to still use WKB is that it gives you a simple form and much better intuition for what the wavefunction looks like outside the transition regions. Moreover, we can derive a generic quantization condition without knowing the details of the transition region, as long as the potential can be approximated as linear there, which we do below.

Explicit Solution for Linear Potential in Transition Regions, Quantization via WKB

To obtain a more explicit solution for the WKB wavefunction and the resulting quantization condition on the energy, we need to explicitly model the transition region wavefunction. We assume a linear approximation to the potential is sufficient. This is certainly consistent with the spirit of the semiclassical approximation. But, if the potential varies wildly in the transition region, one *must* do the full quantum mechanics in that region and one may end up with a different matching condition than we obtain below.

This discussion follows Liboff Section 7.10, which goes into far more detail on WKB than Shankar does.

Derivation for Bound States (cont.)

Let us denote the classical turning points as x_1 and x_2. (Keep in mind that they are dependent on E.) We assume the following form for the solution in the three regions of semiclassical validity:

$$\psi_I(x) = \frac{A}{\sqrt{\kappa(x)}} \exp\left(\frac{1}{\hbar}\int_{x_1}^{x} dx'\kappa(x')\right)$$

$$\psi_{III}(x) = \frac{D}{\sqrt{\kappa(x)}} \exp\left(-\frac{1}{\hbar}\int_{x_2}^{x} dx'\kappa(x')\right)$$

$$\psi_{II,L}(x) = \frac{B}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_{x_1}^{x} dx'p(x') + \delta_L\right)$$

$$\psi_{II,R}(x) = \frac{C}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_{x}^{x_2} dx'p(x') + \delta_R\right)$$

The signs for the arguments in regions *I* and *III* are chosen to ensure the function decays properly (recall, $\kappa(x) \geq 0$ everywhere). For the $_{II,R}$ solution, we shall see that the choice of the order of integration does not matter and our choice is convenient. Note that we reference the integrals to x_1 even though the WKB wavefunctions are not themselves valid for $x \approx x_1$; x_1 is simply a convenient reference point.

Derivation for Bound States (cont.)

We use sines instead of the complex exponential form in the classically allowed region because we proved a long time ago that we can always find real solutions for bound states in one dimension. The parameter freedom is preserved by including both normalization and phase freedom. The two forms of the solution in region *II* are needed in order to do the matching at the left and right sides; they must be consistent, of course, which we shall enforce later. A will be left to be determined by overall normalization, so we will divide everything through by A and work in terms of $b = B/A$, $c = C/A$, and $d = D/A$.

Derivation for Bound States (cont.)

We approximate the potential in the transition regions by

$$V(x) = \begin{cases} E - (x - x_1)\, F_1 & x \approx x_1 \\ E + (x - x_2)\, F_2 & x \approx x_2 \end{cases}$$

We are assured that $V(x_1) = V(x_2) = E$ because we have chosen x_1 and x_2 to be the classical turning points. We know F_1 and F_2 are positive because $E > V(x)$ for $x_1 < x < x_2$. In the transition regions, the Schrödinger Equation takes the form

$$\frac{d^2\psi}{dx^2} + \frac{2\,m\,F_1}{\hbar^2}\,(x - x_1)\,\psi = 0 \qquad \frac{d^2\psi}{dx^2} - \frac{2\,m\,F_2}{\hbar^2}\,(x - x_1)\,\psi = 0$$

These equations are of the same form,

$$\frac{d^2\psi}{dy^2} - y\,\psi = 0 \qquad \text{with} \quad y = \begin{cases} -\left(\frac{2\,m\,F_1}{\hbar^2}\right)^{1/3}(x - x_1) & x \approx x_1 \\ \left(\frac{2\,m\,F_2}{\hbar^2}\right)^{1/3}(x - x_2) & x \approx x_2 \end{cases}$$

One obtains series solutions to the above differential equation in a manner similar to the way we obtained the Hermite polynomials. We will not do that derivation here, but simply quote the results.

Derivation for Bound States (cont.)

The solutions are termed Airy functions. They are given by

$$Ai(y) = a\,f(y) - b\,g(y) \qquad Bi(y) = \sqrt{3}\,[a\,f(y) + b\,g(y)]$$

$$a = \frac{1}{3^{2/3}\,\Gamma\left(\frac{2}{3}\right)} \approx 0.3550 \qquad b = \frac{1}{3^{1/3}\,\Gamma\left(\frac{1}{3}\right)} \approx 0.2588$$

$$f(y) = 1 + \frac{1}{3!}\,y^3 + \frac{1\cdot 4}{6!}\,y^6 + \frac{1\cdot 4\cdot 7}{9!}\,y^9 + \cdots$$

$$g(y) = y + \frac{2}{4!}\,y^4 + \frac{2\cdot 5}{7!}\,y^7 + \frac{2\cdot 5\cdot 8}{10!}\,y^{10} + \cdots$$

where $\Gamma(\,)$ is the Gamma function.

Derivation for Bound States (cont.)

We will need the asymptotic forms ($|y| \gg 1$) of the above functions. From the power series expansions and the factorials in the denominator, it should not be surprising that these asymptotic forms are exponential- or sinusoid-like. They are

$$Ai(y) \overset{y\to+\infty}{\longrightarrow} \frac{1}{2\sqrt{\pi}\, y^{1/4}} \exp\left(-\frac{2}{3} y^{3/2}\right)$$

$$Bi(y) \overset{y\to+\infty}{\longrightarrow} \frac{1}{\sqrt{\pi}\, y^{1/4}} \exp\left(\frac{2}{3} y^{3/2}\right)$$

$$Ai(y) \overset{y\to-\infty}{\longrightarrow} \frac{1}{\sqrt{\pi}\, (-y)^{1/4}} \sin\left(\frac{2}{3} (-y)^{3/2} + \frac{\pi}{4}\right)$$

$$Bi(y) \overset{y\to-\infty}{\longrightarrow} \frac{1}{\sqrt{\pi}\, (-y)^{1/4}} \cos\left(\frac{2}{3} (-y)^{3/2} + \frac{\pi}{4}\right)$$

Clearly, the exponential behavior in one limit and the oscillatory behavior in the other is sensible for matching onto the expoential and oscillatory parts of the WKB wavefunction. In order for the solutions to match onto the decaying exponentials in regions *I* and *III* and the oscillatory solution in region *II*, we need to use only $Ai(y)$. (Recall again that y has the opposite sign as $x - x_1$ when defined near x_1 and the same sign as $x - x_2$ when defined near x_2.) We will use $Bi(y)$ when we consider tunneling or scattering off a positive potential, in which case we need to match onto an oscillatory solution at large distances and a decaying solution inside the potential barrier.

Derivation for Bound States (cont.)

Let us do the matching near x_1. We need to recognize the following

$$p^2(x) \overset{x \to x_1^+}{=} 2\,m\,(E - V(x \to x_1^+)) = 2\,m\,F_1\,(x - x_1) = -\,(2\,m\,F_1\,\hbar)^{2/3}\,y$$

$$\kappa^2(x) \overset{x \to x_1^-}{=} 2\,m\left(V(x \to x_1^-) - E\right) = (2\,m\,F_1\,\hbar)^{2/3}\,y$$

$$2\,m\,F_1\,dx = -\,(2\,m\,F_1\,\hbar)^{2/3}\,dy$$

(Recall that y carries the opposite sign of $x - x_1$.) These forms let us explicitly do the WKB wavefunction argument integrals:

$$\frac{1}{\hbar}\int_{x_1}^{x} dx'\,\kappa(x') = -\int_0^y dy'\,\sqrt{y'} = -\frac{2}{3}\,y^{3/2} \qquad x < x_1, y > 0$$

$$\frac{1}{\hbar}\int_{x_1}^{x} dx'\,p(x') = -\int_0^y dy'\,\sqrt{-y'} = \frac{2}{3}\,(-y)^{3/2} \qquad x > x_1, y < 0$$

Recall that we reference the integrals to x_1 even though the WKB wavefunctions are not themselves valid for $x \approx x_1$; x_1 is simply a convenient reference point.

Derivation for Bound States (cont.)

So the WKB wavefunctions near x_1 become, when written in terms of y,

$$\psi_I(y) = \frac{1}{y^{1/4}} \exp\left(-\frac{2}{3} y^{3/2}\right) \qquad y > 0 \ (x < x_1)$$

$$\psi_{II,L}(y) = \frac{B}{(-y)^{1/4}} \sin\left(\frac{2}{3}(-y)^{3/2} + \delta_L\right) \qquad y < 0 \ (x > x_1)$$

Recall that we already concluded we should only use $Ai(y)$ for $x \approx x_1$. To match onto $Ai(y)$, we must take $B = 2$ and $\delta_L = \pi/4$. $B = 2$ comes from the factor of $1/2$ that appears in the $y \to +\infty$ but not in the $y \to -\infty$ asymptotic versions of $Ai(y)$. Note that the $\frac{1}{\sqrt{\pi}}$ factor in the Airy function asymptotic forms cancels away in connecting $\psi_I(y)$to $\psi_{II,L}(y)$, as does the factor of $(2\,m\,F_1\,\hbar)^{1/6}$ that is the constant of proportionality between $\sqrt{\kappa(x)}$ or $\sqrt{p(x)}$ and $y^{1/4}$.

Repeating the above procedure at $x \approx x_2$ will yield $C = 2D$ and $\delta_R = \pi/4$.

Derivation for Bound States (cont.)

The remaining constant D is determined by matching the $\psi_{II,L}$ and $\psi_{II,R}$ solutions; explicitly:

$$\psi_{II,L}(x) = \psi_{II,R}(x)$$

$$\frac{2}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_{x_1}^{x} dx' p(x') + \frac{\pi}{4}\right) = \frac{2D}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_{x}^{x_2} dx' p(x') + \frac{\pi}{4}\right)$$

Since D must be a constant, the equality can only hold if

$$\frac{1}{\hbar}\int_{x_1}^{x} dx' p(x') + \frac{\pi}{4} = -\left(\frac{1}{\hbar}\int_{x}^{x_2} dx' p(x') + \frac{\pi}{4}\right) + n\pi$$

where n is any integer. We obviously inserted the negative sign for the argument of the sine on the right side so we could combine the two integrals; that is ok to do because D provides the necessary sign freedom. Also, n can be odd, not just even, because $\sin(x + n\pi) = -\sin(x)$ for n odd and the extra minus sign can again be absorbed into D.

Derivation for Bound States (cont.)

Combining the two sides, we obtain

$$\int_{x_1}^{x_2} dx\, p(x) = \left(n\,\pi - \frac{\pi}{2}\right)\hbar = \left(n - \frac{1}{2}\right)\frac{h}{2}$$

That is, the integral of $p(x)$ between the classical turning points must be a half-integral multiple of $h/2$. This is known as the Bohr-Sommerfeld quantization condition

The above condition results in quantization of the energies because the only free parameter in $p(x)$ is E and the above condition puts a constraint on E. We see that energy quantization arises in the same way we described in Section 5.3: it arises from the enforcement of matching conditions, and the reduced freedom in the WKB solution when the state is bound reduces the number of available parameters so that the matching conditions place a condition on E.

Derivation for Bound States (cont.)

It is interesting to rewrite the condition in a different manner to connect it to classical mechanics more explicitly. Classically, the motion is bounded by the turning points. Moreover, the motion must be periodic because the momentum can only take on two values for any given position, corresponding to rightgoing and leftgoing; if the motion were nonperiodic, then p should be allowed to take on an infinite number of values for any x. We are therefore motivated to turn the integral into a line integral over one period of the motion, corresponging to an integral from x_1 to x_2 and back; this gives a factor of 2 because the return path gives an equal contribution as the outgoing path. So,we have

$$\oint dx\, p(x) = \left(n + \frac{1}{2}\right) h$$

(The sign change on the $1/2$ term is absorbed by n).

Derivation for Bound States (cont.)

This integral is the area enclosed in classical phase space ($x - p$ space) by the classical path that the particle would follow. In classical mechanics, this is a very special quantity because it is conserved over time for energy-conserving systems. Moreover, we saw in connection with the unbound WKB solution that this integral is literally the classical action along the classical path. So what we have here is the classical action for one period of the particle's motion in the potential well! The Bohr-Sommerfeld quantization condition thus says that the phase-space area enclosed by the orbit, or equivalently the classical action for one complete period, must be a multiple of h. We thus have a relation that is similar in spirit to the one that showed us that the complex argument of an unbound wavefunction is the classical action as a function of position.

Another way of writing the above relation is to convert from momentum to wavelength, $\lambda(x) = 2\pi\hbar/p(x) = h/p(x)$, which gives

$$\oint \frac{dx}{\lambda(x)} = n + \frac{1}{2}$$

The integral thus integrates up the fractional wavelength accumulated over the orbit. Thus, the total number of wavelengths in one period of the orbit (from x_1 to x_2 and back) must be half-integral.

Section 9

Variational Method

Lecture 27:
The Variational Method

Date Given: 2008/12/03
Date Revised: 2008/12/03

Motivation

It is not always possible to find an analytic solution to the Schrödinger Equation. One can always solve the equation numerically, but this is not necessarily the best way to go; one may not be interested in the detailed eigenfunctions, but rather only in the energy levels and the qualitative features of the eigenfunctions. And numerical solutions are usually less intuitively understandable.

Fortunately, one can show that the values of the energy levels are only mildly sensitive to the deviation of the wavefunction from its true form, and so the expectation value of the energy for an approximate wavefunction can be a very good estimate of the corresponding energy eigenvalue. By using an approximate wavefunction that depends on some small set of parameters and minimizing its energy with respect to the parameters, one makes such energy estimates. The technique is called the variational method because of this minimization process.

This technique is most effective when trying to determine ground state energies, so it serves as a nice complement to the WKB approximation, which works best when one is interested in relatively highly excited states, ones whose deBroglie wavelength is short compared to the distance scale on which the wavelength changes.

The Hamiltonian's Eigenstates and Eigenvalues are Stationary Points of the Energy Functional

For any wavefunction ψ, we may calculate the expectation value of the energy,

$$E[\psi] = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\int dx\, \psi^*(x) H \psi(x)}{\int dx\, \psi^*(x)\psi(x)} \tag{9.1}$$

(We will explain below why we need to explicitly include the normalizing denominator.) We call E a *functional* of the wavefunction ψ because the above expression maps a wavefunction ψ to a number E. We aim to show that we can obtain the energy eigenvalues by requiring that $E[\psi]$ be *stationary* with respect to ψ. By this, we mean that, if there is a function ψ such that, for small variations $\delta\psi$ away from ψ, the corresponding variation δE in $E[\psi]$ vanishes, then ψ is an eigenstate of the Hamiltonian with energy $E[\psi]$. This is the same kind of requirement one places on the classical action in Lagrangian mechanics to yield a differential equation for the classical path, so we are essentially just doing the calculus of variations with the $E[\psi(x)]$ functional instead of the $S[x(t)]$ functional.

Derivation (cont.)

Let's explicitly insert a variation $\delta\psi$ into our equation for $E[\psi]$ and determine the resulting δE:

$$E + \delta E = \frac{\int dx\, [\psi(x) + \delta\psi(x)]^*\, H\, [\psi(x) + \delta\psi(x)]}{\int dx\, [\psi(x) + \delta\psi(x)]^*\, [\psi(x) + \delta\psi(x)]} \tag{9.2}$$

It is now clear why we had to keep the denominator explicit: since we are varying ψ, the normalization of ψ will change. We thus must explicitly include the normalization correction in the denominator to get the correct energy functional. Since having the variations in the denominator on the right side is hard to deal with, but we may multiply through by the denominator to obtain

$$[E + \delta E] \left[\int dx\, [\psi(x) + \delta\psi(x)]^*\, [\psi(x) + \delta\psi(x)] \right] \tag{9.3}$$

$$= \int [\psi(x) + \delta\psi(x)]^*\, H\, [\psi(x) + \delta\psi(x)]$$

$$[E + \delta E] \left[\int dx\, \psi^*(x)\psi(x) + \int dx\, \delta\psi^*(x)\, \psi(x) + \int dx\, \psi^*(x)\, \delta\psi(x) \right] \tag{9.4}$$

$$= \int dx\, \psi^*(x) H\, \psi(x) + \int dx\, \delta\psi^*(x)\, H\, \psi(x) + \int dx\, \psi^*(x) H\, \delta\psi(x)$$

where we have kept only terms to first order in $\delta\psi$.

Derivation (cont.)

Next, we use the unvaried version of the equation to eliminate the terms that include no $\delta\psi$ factors, and we set $\delta E = 0$ to impose the stationarity condition.

$$E\left[\int dx\,\delta\psi^*(x)\,\psi(x) + \int dx\,\psi^*(x)\,\delta\psi(x)\right] = \int \delta\psi^*(x) H\,\psi(x) + \int \psi^*(x) H\,\delta\psi(x) \tag{9.5}$$

$$\int dx\,\delta\psi^*(x)\,(H-E)\,\psi(x) + \int dx\,\psi^*(x)\,(H-E)\,\delta\psi(x) = 0 \tag{9.6}$$

Next, we can show that the two terms must vanish independently by considering two special cases for the variation $\delta\psi$. Suppose $\delta\psi(x) = \chi(x)$ where $\chi(x)$ is purely real. Then we have

$$\int dx\,\chi(x)\,(H-E)\,\psi(x) = -\int dx\,\psi^*(x)\,(H-E)\,\chi(x) \tag{9.7}$$

Next, suppose $\delta\psi$ is completely imaginary, so $\delta\psi(x) = i\,\chi(x)$ where $\chi(x)$ is again real. This gives

$$\int dx\,\chi(x)\,(H-E)\,\psi(x) = \int dx\,\psi^*(x)\,(H-E)\,\chi(x) \tag{9.8}$$

where we have divided out by i.

Derivation (cont.)

Since $\delta\psi$ is completely arbitrary, it is necessary for both equations to hold simultaneously for any real function χ. That is only possible if both terms in the original equation vanish independently. So we have

$$\int dx\, \delta\psi^*(x)\,(H-E)\,\psi(x) = 0 \tag{9.9}$$

$$\int dx\, \psi^*(x)\,(H-E)\,\delta\psi(x) = 0 \tag{9.10}$$

Since $\delta\psi$ is arbitrary in each equation, the integrands must vanish. The first equation yields

$$(H-E)\,\psi(x) = 0 \tag{9.11}$$

One must use the Hermiticity of H to transform the second equation so that H acts on ψ and not $\delta\psi$, which will yield the complex conjugate of the above equation. We recover the eigenvector-eigenvalue equation for H.

Derivation (cont.)

Thus, we have proven that, if $\psi(x)$ is an eigenfunction of H, then a small variation $\delta\psi(x)$ results in no change in E to first order in $\delta\psi(x)$. This proof is interesting in its own right – we had no reason to expect that the Hamiltonian's eigenfunctions would result in an extremum of the energy functional – in much the same way that there is no reason to expect ahead of time that the classical path that solves Newton's equations would correspond to an extremum of the action functional.

We note that the we could have dealt with the normalization of ψ differently. We could have imposed a normalization requirement via a Lagrange multiplier by considering instead the alternate functional

$$E'[\psi] = \int dx\, \psi^*(x)\, H\, \psi(x) - \lambda \left[\int dx\, \psi^*(x)\, \psi(x) - 1 \right] \tag{9.12}$$

and requiring $\delta E' = 0$ under a variation $\delta\psi$. The two functionals $E[\psi]$ and $E'[\psi]$ are the same when one requires that ψ be normalized, so the requirement $\delta E' = 0$ is equivalent to the requirement $\delta E = 0$ under that condition. The result would have been the same had we gone this route.

Derivation (cont.)

Is the converse of what we have proven true? That is, if ψ_m is an eigenfunction of H with eigenvalue E_m, does it hold that the energy functional $E[\psi]$ is stationary with respect to variations away from ψ_m? The answer is yes, and this is relatively easy to show. Let us allow for a variation $\delta\psi$, and let's expand $\delta\psi$ in terms of the eigenstates:

$$|\delta\psi\rangle = \sum_n c_n |\psi_n\rangle \tag{9.13}$$

Now, let's calculate the energy of the wavefunction with the variation:

$$E[\psi_m + \delta\psi] = \frac{(\langle\psi_m| + \langle\delta\psi|)\, H\, (|\psi_m\rangle + |\delta\psi\rangle)}{(\langle\psi_m| + \langle\delta\psi|)\,(|\psi_m\rangle + |\delta\psi\rangle)} \tag{9.14}$$

$$= \frac{E_m + E_m\,(\langle\psi_m|\delta\psi\rangle + \langle\delta\psi|\psi_m\rangle)}{1 + (\langle\psi_m|\delta\psi\rangle + \langle\delta\psi|\psi_m\rangle)} \tag{9.15}$$

$$= E_m \left(\frac{1 + c_m + c_m^*}{1 + c_m + c_m^*}\right) = E_m \tag{9.16}$$

Hence, $\delta E = E[\psi_m + \delta\psi] - E[\psi_m] = 0$. So, indeed, if ψ_m is an eigenstate of H, then $E[\psi]$ is stationary at $\psi = \psi_m$: a variation $\delta\psi$ in ψ_m results in no variation in E to first order in $\delta\psi$.

Derivation (cont.)

For our purposes, the practical implication of the relationship between ψ being an eigenstate and the energy functional being stationary with respect to variations in ψ is that the fractional error in the energy estimate obtained from the trial wavefunction will be much smaller than that fractional error in the wavefunction itself.

Ground State Energy Upper Bounds

It is easy to show that whatever estimate we make for the ground state energy using this technique, it is always an upper bound. Suppose $\psi(x)$ is our trial wavefunction, the energy eigenfunctions are $\{\phi_n(x)\}$ with eigenvalues E_n. We may expand ψ, $\psi(x) = \sum_n c_n \phi_n(x)$. Let us then calculate the energy:

$$E[\psi] = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2} \tag{9.17}$$

Subtract off the ground state energy E_0

$$E[\psi] - E_0 = \frac{\sum_n |c_n|^2 (E_n - E_0)}{\sum_n |c_n|^2} \tag{9.18}$$

Because E_0 is the ground state energy, $E_n - E_0 > 0$ for all n. $|c_n|^2 \geq 0$, so the right side is nonnegative. Therefore

$$E[\psi] \geq E_0 \tag{9.19}$$

Derivation (cont.)

Excited State Energy Estimates

If one knows the eigenfunctions for $n < m$ and one wants to estimate E_m, then one is assured of an upper bound by requiring that the trial wavefunction be orthogonal to the ϕ_n for $n < m$. Explicitly, we consider a trial wavefunction ψ and require

$$\langle \phi_n | \psi \rangle = 0 \quad \text{for} \quad n < m \tag{9.20}$$

This condition can be met for any trial wavefunction $\psi(x)$ via Gram-Schmidt orthogonalization (Section 3.3). We then calculate the energy functional

$$E[\psi] = \frac{\sum_{n=m}^{\infty} |c_n|^2 E_n}{\sum_{n=0}^{\infty} |c_n|^2} \tag{9.21}$$

because the terms with $n < m$ vanish due to $c_n = \langle \phi_n | \psi \rangle = 0$. We then subtract E_m:

$$E[\psi] - E_m = \frac{\sum_{n=m}^{\infty} |c_n|^2 (E_n - E_m)}{\sum_{n=0}^{\infty} |c_n|^2} \tag{9.22}$$

$E_n \geq E_m$ for $n > m$, so we are assured the right side is nonnegative, yielding

$$E[\psi] \geq E_m \quad \text{when} \quad \langle \phi_n | \psi \rangle = 0 \quad \text{for} \quad n < m \tag{9.23}$$

Derivation (cont.)

Of course, the above is somewhat useless because, if one is able to exactly solve the eigenvector-eigenvalue equation for some n, then one is usually able to solve it for all n and the above technique is unnecessary. The idea about Gram-Schmidt orthogonalization is good, though: even if one only has approximations to the lower-energy state eigenfunctions, one should still construct trial wavefunctions for the higher states using Gram-Schmidt orthogonalization. One can show, for example, that the error in estimating E_m is related to the mismatch between one's estimates ψ_n for the actual eigenfunctions ϕ_n for $n < m$. In particular, if

$$\delta_0 = 1 - |\langle \psi_0 | \phi_0 \rangle|^2 \tag{9.24}$$

expresses the fractional deviation of the approximate ground state wavefunction ψ_0 from the true one ϕ_0, then one can show that a trial wavefunction ψ_1 that has been constructed using Gram-Schmidt orthogonalization with respect to ψ_0 (not ϕ_0!) yields

$$E[\psi_1] \geq E_1 - \delta_0 (E_1 - E_0) \tag{9.25}$$

That is, $E[\psi_1]$ is no longer an upper bound on E_1, but the amount by which it underestimates E_1 is proportional to δ_0. Because δ_0 is quadratic in $\langle \psi_0 | \phi_0 \rangle$, the fractional error in estimating E_1 is much smaller than the error in estimating the wavefunction, as long as $E_1 - E_0$ is of order E_1.

How it Works in Practice

How do we construct trial wavefunctions for the purpose of making these variational estimates of energy eigenvalues? One cannot usually guess the form for the correct wavefunction exactly. But one usually knows the general features of the wavefunction one wants. So one constructs a trial wavefunction that depends on some small set of parameters, calculates the energy functional for the trial wavefunction as a function of this small set of parameters, and then requires E be stationary with respect to those parameters – *i.e.*, that all the partial derivatives of E with respect to the parameters vanish. In effect, we are explicitly applying the stationarity condition to some subset of all possible wavefunctions under the expectation that we can get a very good approximation to the energy with a good approximation of the correct wavefunction thanks to the stationarity of E with respect to variations in ψ.

An important guide is to require that the trial wavefunctions be eigenfunctions of any other Hermitian operators A that commute with H. These are usually related to symmetries of the Hamiltonian. For example, a Hamiltonian in which the potential is an even function of posiition commutes with the parity opertator and hence any eigenfunctions of H must have definite parity. The generators of continuous symmetry transformations are also good examples; rotation symmetry, for example, implies that angular momentum commutes with H. If one makes these requirements on the trial wavefunctions, one is assured that they are at least members of the same subspace as the true eigenfunctions.

Applications (cont.)

Example 9.1: Particle in a Box

A good first example is to do a variational-method estimate for a problem whose exact solution we already know – the particle in a box. The Hamiltonian guides our choice of trial wavefunction in two ways. First, we know the trial wavefunction should vanish at the box edge and outside. If the wavefunction's derivative is large at any point, the kinetic energy will be big, so the wavefunction should rise smoothly and slowly away from zero at the box edges. Second, the Hamiltonian commutes with the parity operator, so the trial wavefunction should be even or odd. We try an even function because it can have no zeroes and thus also have the smallest possible derivative and hence kinetic energy. We try

$$\psi(x;c) = \begin{cases} \left(\left(\frac{L}{2}\right)^2 - x^2\right)(1 + c\,x^2) & |x| \leq \frac{L}{2} \\ 0 & |x| > \frac{L}{2} \end{cases} \tag{9.26}$$

This is the simplest polynomial trial function we can use. Evenness requires that we only include even powers of x. A simple quadratic could satisfy the requirement of going to zero at the box edges, but admits no free parameters: it must be $\left((L/2)^2 - x^2\right)$ (an overall constant multiplier just sets the normalization and cancels out between the numerator and denominator of the energy functional). So the next possibility is a polynomial containing x^2 and x^4, and the above is just one convenient way to parameterize it (again, neglecting any overall scale factor).

Applications (cont.)

Calculating the integrals for the energy functional is a somewhat nontrivial algebraic exercise that can be done correctly in Mathematica, yielding

$$E[\psi] \to E[c] = \frac{3\,\hbar^2}{m\,L^2}\,\frac{11\left(\frac{L}{2}\right)^4 c^2 + 14\left(\frac{L}{2}\right)^2 c + 35}{\left(\frac{L}{2}\right)^4 c^2 + 6\left(\frac{L}{2}\right)^2 c + 21} \tag{9.27}$$

We then find the extrema of $E(c)$ with respect to c, $dE/dc = 0$, yielding

$$26\left(\frac{L}{2}\right)^4 c^2 + 196\left(\frac{L}{2}\right)^2 c + 42 = 0 \tag{9.28}$$

with solutions

$$c_1 = -\frac{0.221075}{\left(\frac{L}{2}\right)^2} \qquad c_2 = -\frac{7.31771}{\left(\frac{L}{2}\right)^2} \tag{9.29}$$

Applications (cont.)

We may then calculate E for these values, giving

$$E(c_1) = 4.93488\,\frac{\hbar^2}{m\,L^2} \qquad E(c_2) = 51.0652\,\frac{\hbar^2}{m\,L^2} \tag{9.30}$$

The reason that the second solution has so much higher energy is that it has zeros at $x = \pm c_2^{-1/2} \approx 0.185\,L$, hence its derivative and its kinetic energy is much larger than the c_1 state, which has no zeros. The true ground state energy is

$$E_{n=1} = \frac{\hbar^2\pi^2}{2\,m\,L^2} = 4.93480\,\frac{\hbar^2}{m\,L^2} \tag{9.31}$$

which is shockingly close to $E(c_1)$. Plotting the trial function shows that it is a very good match to the true ground state wavefunction.

Applications (cont.)

Another interesting fact is that the second solution, using c_2, is a decent approximation to the second excited state, $n = 3$:

$$E_{n=3} = 9\, E_{n=1} = 44.4132\, \frac{\hbar^2}{m\, L^2} \tag{9.32}$$

There is a general theorem on this point, the Hylleraas-Undheim Theorem, which essentially states that, if the trial wavefunction depends on a set of parameters, the alternate solutions for the parameters giving non-minimal extrema of E yield estimates for the energies of the excited states of the system.

Example 9.2: Other Examples

A few other examples of simple analytically soluble problems that can be approximated quite well or perfecty by the variational method are:

- The simple harmonic oscillator – a Gaussian trial function will recover the ground state energy and wavefunction exactly because the ground state wavefunction is a Gaussian. A Gaussian times a polynomial will cover the excited states exactly, also, via the Gram-Schmidt orthogonalization procedure.
- Bound state of a δ-function potential well. The trial function should be a decaying exponential.

Section 10

Classical Limit

Lecture 28:
Classical Limit

Date Given: 2008/12/05
Date Revised: 2008/12/05

Ehrenfest's Theorem

How do expectation values evolve in time? We expect that, as quantum effects becomes small, the fractional uncertainty in a physical observable Ω, given by $\sqrt{\langle(\Delta\Omega)^2\rangle}/\langle\Omega\rangle$, for a state $|\psi\rangle$, becomes small and we need only consider the expectation value $\langle\Omega\rangle$, not the full state $|\psi\rangle$. So, the evolution of expectation values should approach classical equations of motion as $\hbar \to 0$. To check this, we must first calculate how expectation values time-evolve, which is easy to do:

$$\begin{aligned}
\frac{d}{dt}\langle\Omega\rangle &= \left(\frac{d}{dt}\langle\psi|\right)\Omega|\psi\rangle + \langle\psi|\Omega\left(\frac{d}{dt}|\psi\rangle\right) + \langle\psi|\left(\frac{d\Omega}{dt}\right)|\psi\rangle \\
&= -\frac{i}{\hbar}\left[-\left(\langle\psi|H\right)\Omega|\psi\rangle + \langle\psi|\Omega\left(H|\psi\rangle\right)\right] + \langle\psi|\left(\frac{d\Omega}{dt}\right)|\psi\rangle \\
&= -\frac{i}{\hbar}\langle\psi|[\Omega, H]|\psi\rangle + \langle\psi|\left(\frac{d\Omega}{dt}\right)|\psi\rangle \\
&= -\frac{i}{\hbar}\langle[\Omega, H]\rangle + \left\langle\frac{d\Omega}{dt}\right\rangle
\end{aligned} \tag{10.1}$$

Ehrenfest's Theorem (cont.)

If the observable Ω has no explicit time-dependence, this reduces to

$$\frac{d}{dt}\langle\Omega\rangle = -\frac{i}{\hbar}\langle[\Omega, H]\rangle \tag{10.2}$$

Equations 10.1 and 10.2 are known as the Ehrenfest Theorem.

Ehrenfest's Theorem (cont.)

Applications of the Ehrenfest Theorem

To make use of it, let's consider some examples. For $\Omega = P$, we have

$$\frac{d}{dt}\langle P\rangle = -\frac{i}{\hbar}\langle [P, H]\rangle$$

We know P commutes with $P^2/2\,m$, so the only interesting term will be $[P, V(X)]$. We can see what that is via Taylor expansion of $V(X)$:

$$\begin{aligned}[P, V(X)] &= \left[P, \sum_{n=0}^{\infty} V_n X^n\right] = \sum_{n=0}^{\infty} V_n\,[P, X^n] \\ &= \sum_{n=0}^{\infty} V_n\,\left(-i\,\hbar\,n\,X^{n-1}\right) = -i\,\hbar\,\frac{dV(X)}{dX}\end{aligned}$$

where $[P, X^n] = -i\,\hbar\,n\,X^{n-1}$ can easily be proven via induction and $\frac{dV(X)}{dX}$ should be interpreted as "take the derivative of $V(x)$ and replace x with X everywhere."

Ehrenfest's Theorem (cont.)

So we have

$$\frac{d}{dt}\langle P\rangle = -\left\langle \frac{dV}{dX}\right\rangle$$

which is the standard classical result for Newton's second law in the presence of a potential $V(x)$. Note that this result holds for arbitrary $\hbar$; it only becomes a good description of the state, though, when $\langle(\Delta X)^2\rangle/\langle X\rangle^2 \ll 1$ and $\langle(\Delta P)^2\rangle/\langle P\rangle^2 \ll 1$.

Another classical result is found by setting $\Omega = X$:

$$\frac{d}{dt}\langle X\rangle = -\frac{i}{\hbar}\left\langle [X, \frac{P^2}{2\,m}]\right\rangle$$

By the same kind of calculation as applied above, $[X, P^2] = 2\,i\,\hbar\,P$, so

$$\frac{d}{dt}\langle X\rangle = \frac{\langle P\rangle}{m}$$

which is the classical relation between velocity and linear momentum.

Ehrenfest's Theorem (cont.)

The above two results can be rewritten as

$$\frac{d}{dt}\langle P\rangle = -\left\langle \frac{dH}{dX} \right\rangle \qquad \frac{d}{dt}\langle X\rangle = \left\langle \frac{dH}{dP} \right\rangle$$

because, for the first equation, the derivative of the kinetic term in H with respect to X vanishes, and, for the second equation, the derivative of the potential term with respect to P vanishes. These we recognize as Hamilton's equations.

Poisson Brackets and Commutators

Recall in classical mechanics that the Poisson Bracket of two functions $f(x,p)$ and $g(x,p)$ of the the state variables x and p is

$$\{f,g\}_{q,p} = \frac{\partial f}{\partial x}\frac{\partial g}{\partial p} - \frac{\partial g}{\partial x}\frac{\partial f}{\partial p}$$

We have in classical mechanics the generic result for the time evolution of a function F of the state variables x, p

$$\frac{df}{dt} = \{f,\mathcal{H}\} + \frac{\partial f}{\partial t}$$

where $\mathcal{H}(x,p)$ is the classical Hamiltonian. Notice how similar this is to Ehrenfest's Theorem; if one replaces $f(x,p)$ with the expectation value of the analogous quantum operator $F(X,P)$ and the Poisson bracket with $-\frac{i}{\hbar}[F,H]$, one obtains Ehrenfest's Theorem.

Correspondences between Classical and Quantum Mechanics (cont.)

The above general classical expression for time evolution of a function $f(x,p)$ yields Hamilton's equations when applied to $f = x$ and $f = p$:

$$\frac{dx}{dt} = \{x, \mathcal{H}\} = \frac{\partial \mathcal{H}}{\partial p} \qquad \frac{dp}{dt} = \{p, \mathcal{H}\} = -\frac{\partial \mathcal{H}}{\partial x}$$

We saw earlier the quantum version of these equations, again with $\{\ ,\ \}$ replaced by $-\frac{i}{\hbar}[\ ,\ ]$.

Now, recall that Postulate 2 can be rewritten to specify $[X,P]$ rather than the matrix elements of X and P in the $\{|x\rangle\}$ basis. We showed that the two formulations were equivalent. But the latter formulation was a postulate, not derivable in any way. The analogies we've drawn between classical and quantum mechanics suggest a prescription that leads from classical mechanics to $[X,P] = i\hbar$. Specifically, we know in classical mechanics that $\{x,p\}_{x,p} = 1$ (just calculate it). We've seen above that, in many cases, quantum mechanical Ehrenfest relations reduce to the equations of classical mechanics if we make the correspondence $-\frac{i}{\hbar}[\ ,\ ] \leftrightarrow \{\ ,\ \}_{q,p}$ where q and p are the variables that define the classical trajectory and that are replaced by the operators Q and P in quantum mechanics. This correspondence would give, for example, $[X,P] = i\hbar$. Morover, it provides a mean to convert any classical problem to a quantum mechanical problem. It's not a derivation, but it is a prescription for what the Postulate 2 commutators should be; along with the other postulates, it defines quantum mechanics completely.

Correspondences between Classical and Quantum Mechanics (cont.)

The next interesting correspondence is a direct one between the Schrödinger Equation and the Hamilton-Jacobi equation of classical mechanics (don't worry if you don't know know what the latter is.)

You of course know and love the Schrödinger Equation, which, projected onto the $\{|x\rangle\}$ basis, is

$$-\frac{\hbar^2}{2\,m}\,\frac{\partial^2\psi}{\partial x^2} + V(x)\,\psi = i\,\hbar\,\frac{\partial\psi}{\partial t}$$

where $V(x)$ is the classical potential in which the particle moves. We know from the WKB approximation that a reasonable approximate form for the wavefunction is

$$\psi(x,t) = \sqrt{\rho(x,t)}\,e^{\frac{i}{\hbar}\,S(x,t)} \tag{10.3}$$

Correspondences between Classical and Quantum Mechanics (cont.)

Let's insert the above form into the Schrödinger Equation:

$$-\frac{\hbar^2}{2\,m}\left\{\frac{\partial^2\sqrt{\rho}}{\partial x^2}+\frac{2\,i}{\hbar}\,\frac{\partial\sqrt{\rho}}{\partial x}\,\frac{\partial S}{\partial x}-\frac{1}{\hbar^2}\,\sqrt{\rho}\left(\frac{\partial S}{\partial x}\right)^2+\frac{i}{\hbar}\,\sqrt{\rho}\,\frac{\partial^2 S}{\partial x^2}\right\}e^{\frac{i}{\hbar}S}+V(x)\,\sqrt{\rho}\,e^{\frac{i}{\hbar}S}$$
$$=i\,\hbar\left\{\frac{\partial\sqrt{\rho}}{\partial t}+\frac{i}{\hbar}\,\sqrt{\rho}\,\frac{\partial S}{\partial t}\right\}e^{\frac{i}{\hbar}S}$$

The equation contains terms up to second order in $\hbar$; neglect all terms containing $\hbar$ or $\hbar^2$. (To do this correctly, one ought to really divide through by $S(x)$ and discard all terms that have the dimensionless parameter $\hbar/S(x) \ll 1$, but that will yield the same result.) This leaves (canceling the common $\sqrt{\rho}\,e^{\frac{i}{\hbar}S}$)

$$\frac{1}{2\,m}\left(\frac{\partial S}{\partial x}\right)^2+V(x)+\frac{\partial S}{\partial t}=0$$

This is the Hamilton-Jacobi Equation of classical mechanics. It is equivalent to (in fact, derived from) Hamilton's Equations. It is not particularly remarkable that we obtain a classical equation of motion by taking the classical limit of the Schrödinger Equation; we've done the same thing with the Ehrenfest relations. The remarkable thing is the role the classical action plays in the quantum-mechanical problem in the semiclassical limit, as we saw in connection to the WKB approximation.

Section 11

Multiparticle Systems

Lecture 29:
Multiparticle Systems
Direct Product Spaces: States

Date Revised: 2009/01/08
Date Given: 2009/01/05

Direct Product Spaces

Suppose we have two Hilbert spaces, $\mathbb{V}_1$ and $\mathbb{V}_2$, each containing the states corresponding to a particular degree of freedom (dof); a typical example is that $\mathbb{V}_1$ contains the states for particle 1 and $\mathbb{V}_2$ for particle 2, where both particles live in a single spatial dimension. Then we can build a new Hilbert space, $\mathbb{V} = \mathbb{V}_1 \otimes \mathbb{V}_2$, that contains the state of the two particles considered together. This space is called a direct product space. Formally, we construct the elements of the space in three steps:

- First, we define elements that are combinations of single states from the two factor spaces:

$$|v,\, w\rangle^{(1)\otimes(2)} \equiv |v\rangle^{(1)} \otimes |w\rangle^{(2)}$$

  where the superscripts on each ket outside the ket bracket indicates which particle's Hilbert space it belongs to: the $^{(1)}$ kets belong to $\mathbb{V}_1$, the $^{(2)}$ kets to $\mathbb{V}_2$ and the $^{(1)\otimes(2)}$ kets to $\mathbb{V} = \mathbb{V}_1 \otimes \mathbb{V}_2$. We emphasize that this definition cannot be algebraically reduced to something simpler. An example of the above would be for $|v\rangle^{(1)}$ to be a basis element of the position-basis representation for particle 1, $|\widehat{x}\rangle$, and for $|w\rangle^{(2)}$ to be a basis element of the position-basis element for particle 2, $|\widetilde{x}\rangle$. The direct product vector

Direct Product Spaces (cont.)

$|\widehat{x}, \widetilde{x}\,\rangle^{(1)\otimes(2)} = |\widehat{x}\,\rangle^{(1)} \otimes |\widetilde{x}\,\rangle^{(2)}$ is simply the state in which particle 1 is at $\widehat{x}$ and particle 2 is at $\widetilde{x}$; it can be written in no simpler fashion.

Note that many authors would have written $|x_1\,\rangle$, $|x_2\,\rangle$ and $|x_1, x_2\,\rangle$, dropping the superscripts altogether and relying on context to make it clear which state refers to particle 1, particle 2, and the combination of the two. There is nothing wrong with this, but the use of the numbered subscripts makes it seem that the position value x_1 is only available to the first particle and the position value x_2 only to the second particle. However, the particles live in the same physical space, both x_1 and x_2 are accessible to both, and it would be perfectly reasonable to have $|x_2\,\rangle^{(1)}$, $|x_1\,\rangle^{(2)}$, and $|x_2, x_1\,\rangle^{(1)\otimes(2)} = |x_2\,\rangle^{(1)} \otimes |x_1\,\rangle^{(2)}$.

- Second, all possible linear combinations operate in the expected manner, assuming the two factor spaces $\mathbb{V}_1$ and $\mathbb{V}_2$ have the same field (complex numbers for what we usually consider):

$$\Big(\alpha|v\,\rangle^{(1)}\Big) \otimes \Big(\beta|w\,\rangle^{(2)}\Big) = \alpha\beta\,|v, w\,\rangle^{(1)\otimes(2)}$$

$$\Big(\alpha_1|u_1\,\rangle^{(1)} + \alpha_2|v_1\,\rangle^{(1)}\Big) \otimes |w\,\rangle^{(2)} = \alpha_1|u_1, w\,\rangle^{(1)\otimes(2)} + \alpha_2|v_1, w\,\rangle^{(1)\otimes(2)}$$

$$|u\,\rangle^{(1)} \otimes \Big(\beta_1|v_2\,\rangle^{(2)} + \beta_2|w_2\,\rangle^{(2)}\Big) = \beta_1|u, v_2\,\rangle^{(1)\otimes(2)} + \beta_2|u, w_2\,\rangle^{(1)\otimes(2)}$$

Direct Product Spaces (cont.)

- Third, the inner product is defined as the obvious extension of the inner product in each space:

$$ {}^{(1)\otimes(2)}\langle v_1, w_1 \,|\, v_2, w_2 \,\rangle^{(1)\otimes(2)} \equiv \left({}^{(1)}\langle v_1 \,|\, v_2 \,\rangle^{(1)}\right)\left({}^{(2)}\langle w_1 \,|\, w_2 \,\rangle^{(2)}\right) $$

By including linear combinations and requiring the scalar fields of the two products spaces be the same, we ensure that the direct product space is a vector space. A reasonable basis for the product space is simply the set of direct products of the bases of the individual spaces; that is, suppose $\{|n\rangle^{(1)}\}$ are a basis for the first space and $\{|m\rangle^{(2)}\}$ are a basis for the second space. Then a basis for the direct product space consists of all products of the form

$$ |n, m\rangle^{(1)\otimes(2)} = |n\rangle^{(1)} \otimes |m\rangle^{(2)} $$

where both n and m run over their full range of values. If the two factor spaces have dimension N and M (with these possibly being infinite), then the direct product space has dimension $N \times M$.

Direct Product Spaces (cont.)

By defining the inner product as above, the direct product space inherits all the necessary inner product properties from the factor spaces, rendering the direct product space an inner product space.

Finally, the restriction to normalizable states that occurs to render the inner product space into the physical Hilbert space is inherited via the inherited definition of inner product. So we are automatically assured that all elements of the direct product space are normalizable if the factor-spaces are physical Hilbert spaces.

Note that the fact that the direct product space is a physical Hilbert space ensures that Postulate 1 continues to be satisfied.

The Null Vector, Inverse Vectors, Invertibility, and Entanglement

The null vector and inverse vectors are a bit tricky in direct product spaces because there are multiple ways to construct them. First, any direct product in which one factor is a null vector from either space gives the null vector of the direct product space:

$$|0\rangle^{(1)} \otimes |w\rangle^{(2)} = |0\rangle^{(1)\otimes(2)} = |v\rangle^{(1)} \otimes |0\rangle^{(2)}$$

We can see that the two factor forms are equivalent by calculating their norms: in each case, the norm vanishes because the norm of the direct product is the product of the norms, and one factor has vanishing norm in either case. The definition of inner product requires that the null vector be the only vector that has vanishing norm, so we must take as a *definition* that all these ways of obtaining $|0\rangle^{(1)\otimes(2)}$ are equivalent in order for the direct product space to be an inner product (and hence physical Hilbert) space.

An implication of this is that the mapping from the two factor spaces to the direct product space is not *one-to-one* and hence is *noninvertible*.

Direct Product Spaces (cont.)

The same issue arises for inverses. There are multiple pairs in the factor spaces that map to the inverse of a given member of the direct product space:

$$-|v,w\rangle^{(1)\otimes(2)} = -\left(|v\rangle^{(1)} \otimes |w\rangle^{(2)}\right)$$
$$= \left(-|v\rangle^{(1)}\right) \otimes |w\rangle^{(2)} = |v\rangle^{(1)} \otimes \left(-|w\rangle^{(2)}\right)$$

We can see another way in which the mapping from factor space pairs to the direct product space is noninvertible: in addition to not being one-to-one, it is also not *onto*. That is, not every element of the direct product space can be written purely as a product of elements of the factor spaces. A simple example can be construction from any two basis elements of the direct product space:

$$|n_1\rangle^{(1)} \otimes |m_1\rangle^{(2)} + |n_2\rangle^{(1)} \otimes |m_2\rangle^{(2)} \neq |v\rangle^{(1)} \otimes |w\rangle^{(2)}$$
$$\text{for all } |v\rangle^{(1)} \text{ in } \mathbb{V}_1 \text{ and all } |w\rangle^{(2)} \text{ in } \mathbb{V}_2$$

It is easy to prove this by assuming that the above is true, expanding the $|v\rangle^{(1)}$ in terms of the $\{|n\rangle^{(1)}\}$ and $|w\rangle^{(2)}$ in terms of the $\{|m\rangle^{(2)}\}$, and obtaining a contradiction. We will do this later.

Direct Product Spaces (cont.)

States of the above type are called entangled — neither degree of freedom (think "particle") is in a particular state in its own space! This is a fundamentally quantum mechanical phenomenon that arises from the fact that the state of a particle is represented by vectors in a Hilbert space and that these Hilbert spaces can be direct producted together in a noninvertible manner. Entanglement, which arises from the noninvertibility, makes the physics of systems with multiple degrees of freedom more than just the some of the parts.

Direct Products vs. Direct Sums

Recall that we defined direct sums early on in the course in connection with the idea of vector subspaces: a direct sum space $\mathbb{V}_1 \oplus \mathbb{V}_2$ consists of all linear combinations of elements of the two spaces $\mathbb{V}_1$ and $\mathbb{V}_2$. Since it is required that one be able to add elements of $\mathbb{V}_1$ and $\mathbb{V}_2$, they must already be subspaces of a larger vector space $\mathbb{V}$: that is, elements of $\mathbb{V}_1$ and of $\mathbb{V}_2$ already belong to $\mathbb{V}$ and there is already a rule for how to add them.

A direct product space is quite different. Perhaps the most interesting difference is the fact that that the direct product space $\mathbb{V}_1 \otimes \mathbb{V}_2$ requires no prior existence of a vector space $\mathbb{V}$ in which $\mathbb{V}_1 \otimes \mathbb{V}_2$ is contained. That is, $\mathbb{V}_1 \otimes \mathbb{V}_2$ is a **new** vector space that is in no way reducible in terms of $\mathbb{V}_1$ and $\mathbb{V}_2$ separately. Of course, $\mathbb{V}_1 \otimes \mathbb{V}_2$ may look like a space we have already seen in some cases, but that is not a generic statement. In general, $\mathbb{V}_1 \otimes \mathbb{V}_2$ is just a new object.

One specific technical difference between direct sum and direct product spaces is that the construction of the former is invertible while that of the latter is not. Any element of $\mathbb{V}_1 \oplus \mathbb{V}_2$ can be written uniquely as a sum of elements of $\mathbb{V}_1$ and $\mathbb{V}_2$: just decompose it in terms of the basis of $\mathbb{V}_1 \oplus \mathbb{V}_2$ and split up the terms into basis elements belonging to $\mathbb{V}_1$ and to $\mathbb{V}_2$. As we explained above, such a decomposition is not in general possible for elements of a direct product space.

Pitfalls in Understanding Direct Product Spaces

In the case of taking the direct product of the Hilbert spaces of two particles that live in the same spatial space, a typical pitfall is the desire to put the particles in the same Hilbert space. Be wary of this! Except through explicit interactions (which we will encounter later) of the two particles, *the Hilbert space states of the two particles are totally independent* even though, when one projects onto, say, the position basis, both wavefunctions give the probability of detecting the two particles in the same physical space and may overlap.

Another possible stumbling block: don't confuse the Hilbert space dimension with the dimension of the physical space that the particles live in; we will give examples below.

Direct Product Spaces (cont.)

Examples of Direct Product Spaces

- As we have indicated above, one can construct a direct product Hilbert space from multiple single-dof Hilbert spaces. For the free particle, each factor space has dimension equal to the size of the real numbers, so the product space has that size squared (which is the same). For the particle in a box or SHO, the number of single-dof states is countably infinite; the square of that number is also countably infinite. In both cases, the two particles move about in a single, shared spatial dimension, and the number of degrees of spatial freedom is two: the spatial coordinates of the two particles.
- Different spatial degrees of freedom of a single particle can be put together via a direct product to give the full two- or three-spatial-dimensional state of that particle.
- One can combine spatial and other degrees of freedom. For a hydrogen atom, one factor space would consist of the three-spatial-dimensional center-of-mass position, while the other would consist of the three-spatial-dimensional relative electron-proton position (described by the radial quantum number n and the angular momentum quantum numbers l^2 and l_z, which we will cover in detail later). Another example would be a rigid rotator, where again one factor space is the center-of-mass position and the other is the same l^2 and l_z quantum numbers (there is no radial quantum number because the body is rigid).
- One can of course combine multiple degrees of freedom for multiple particles.

Lecture 30:
Multiparticle Systems
Direct Product Spaces:
States, Operators, Hamiltonians, Time Evolution
Separable vs. Non-Separable Hamiltonians

Date Revised: 2009/01/08

Date Given: 2009/01/07

Expansion of Direct Product Space States in Terms of Basis States

Our definition of direct product space states has so far been abstract. Let's pick a basis and expand states in terms of that basis so we can get a feel for how direct products work. These expansions are the foundation for any kind of matrix representation.

Let's use $\{|n_j\rangle^{(1)}\}$ as the basis for $\mathbb{V}_1$ and $\{|m_k\rangle^{(2)}\}$ as the basis for $\mathbb{V}_2$. Let $|v\rangle^{(1)}$ and $|w\rangle^{(2)}$ be vectors in the two spaces. Then we have

$$|v\rangle^{(1)} = \sum_j v_j |n_j\rangle^{(1)} \qquad |w\rangle^{(2)} = \sum_k w_k |m_k\rangle^{(2)}$$

$$\begin{aligned} |v\rangle^{(1)} \otimes |w\rangle^{(2)} &= \left(\sum_j v_j |n_j\rangle^{(1)}\right) \otimes \left(\sum_k w_k |m_k\rangle^{(2)}\right) \\ &= \sum_{j,k} v_j\, w_k \left(|n_j\rangle^{(1)} \otimes |m_k\rangle^{(2)}\right) = \sum_{j,k} v_j\, w_k \left(|n_j, m_k\rangle^{(1)\otimes(2)}\right) \end{aligned}$$

Direct Product Spaces (cont.)

However, the above state is not the most generic possible state in the direct product space; that would be

$$|u\rangle^{(1)\otimes(2)} = \sum_{j,k} u_{jk} \, |n_j, m_k\rangle^{(1)\otimes(2)} \tag{11.1}$$

For arbitrary $\{u_{jk}\}$, one cannot decompose u_{jk} in the form $u_{jk} = v_j w_k$. In particular, let's consider the example from earlier:

$$|u\rangle^{(1)\otimes(2)} = |n_1\rangle^{(1)} \otimes |m_1\rangle^{(2)} + |n_2\rangle^{(1)} \otimes |m_2\rangle^{(2)}$$

We have

$$u_{11} = 1 \qquad u_{12} = 0 \qquad u_{21} = 0 \qquad u_{22} = 1$$

Let's assume an expansion $u_{jk} = v_j w_k$. The statement $u_{12} = 0$ implies that either $v_1 = 0$ or $w_2 = 0$. But then one of $u_{11} = v_1 w_1$ or $u_{22} = v_2 w_2$ must vanish; they do not. Contradiction.

Operators on Direct Product Spaces

We have constructed the vectors belonging to the direct product space. How do the factor-space operators go over to the direct product space?

Recall our definition of operators: an operator Ω is a rule that associates to each state $|v\rangle$ in a Hilbert space another state $|w\rangle = \Omega|v\rangle$. This immediately makes it clear that, when consider a direct product space $\mathbb{V} = \mathbb{V}_1 \otimes \mathbb{V}_2$, an operator $\Omega^{(1)}$ that acts in $\mathbb{V}_1$ simply does not exist in $\mathbb{V}_2$. The natural thing to do, then, is to assign it the action of the identity operator in $\mathbb{V}_2$, $I^{(2)}$. This is an *assumption*: there is no reason that we *must* do it this way, though it certainly seems like the most sensible thing to do. Symbolically, we have

$$\Omega^{(1)\otimes(2)} \equiv \Omega^{(1)} \otimes I^{(2)}$$

or, explicitly,

$$\Omega^{(1)\otimes(2)}|v, w\rangle^{(1)\otimes(2)} = \left(\Omega^{(1)}|v\rangle^{(1)}\right) \otimes \left(I^{(2)}|w\rangle^{(2)}\right)$$

Note that the individual factors in any operator direct product commute to the extent that it does not matter in what order the factor operators and states are written as long as the labeling makes it clear which operators act in which space.

Direct Product Spaces (cont.)

For the sake of brevity, we will suppress the superscripts when it is clear from context what is going on. For example, in considering two one-dimensional particles, it is unambiguous to just write X_1 and X_2. If we consider product states of basis elements in the position-basis representation, we have

$$X_1|\widehat{x}\rangle^{(1)} = \widehat{x}\,|\widehat{x}\rangle^{(1)} \qquad X_2|\widetilde{x}\rangle^{(2)} = \widetilde{x}\,|\widetilde{x}\rangle^{(2)}$$

$$X_1\left(|\widehat{x}\rangle^{(1)} \otimes |\widetilde{x}\rangle^{(2)}\right) = \left(X_1|\widehat{x}\rangle^{(1)}\right) \otimes \left(I\,|\widetilde{x}\rangle^{(2)}\right) = \widehat{x}\left(|\widehat{x}\rangle^{(1)} \otimes |\widetilde{x}\rangle^{(2)}\right)$$

$$X_2\left(|\widehat{x}\rangle^{(1)} \otimes |\widetilde{x}\rangle^{(2)}\right) = \left(I\,|\widehat{x}\rangle^{(1)}\right) \otimes \left(X_2\,|\widetilde{x}\rangle^{(2)}\right) = \widetilde{x}\left(|\widehat{x}\rangle^{(1)} \otimes |\widetilde{x}\rangle^{(2)}\right)$$

The two I operators are different: they are $I^{(1)}$ (last line) and $I^{(2)}$ (second line), but the context makes it clear that the I operator in the second line is $I^{(2)}$ and the I operator in the last line is $I^{(1)}$.

Direct Product Spaces (cont.)

While this prescription for how to construct direct product space operators was fairly straightforward, we also need to test that it is consistent with the postulates of quantum mechanics. The main thing to test is Postulate 2: do the matrix elements (or commutators) still behave as we want them to. The answer is yes, trivially, because of the way the inner product factorizes among the factor spaces. Specifically, let's calculate the matrix elements of X_j and P_j for a direct product space consisting of N factors:

$$\left\langle \prod_{p=1}^{N} x_p \right| X_j^{\prod_{n=1}^{N}(n)} \left| \prod_{q=1}^{N} x_q' \right\rangle = \langle x_j | X_j^{(j)} | x_j' \rangle \prod_{n\neq j}^{N} \langle x_n | I^{(n)} | x_n' \rangle$$

$$= x_j \, \delta(x_j - x_j') \prod_{n\neq j}^{N} \delta(x_n - x_n') = x_j \prod_{n=1}^{N} \delta(x_n - x_n')$$

$$\left\langle \prod_{p=1}^{N} x_p \right| P_j^{\prod_{n=1}^{N}(n)} \left| \prod_{q=1}^{N} x_q' \right\rangle = \langle x_j | P_j^{(j)} | x_j' \rangle \prod_{n\neq j}^{N} \langle x_n | I^{(n)} | x_n' \rangle$$

$$= \left(-i\hbar \frac{d}{dx_j} \delta(x_j - x_j') \right) \prod_{n\neq j}^{N} \delta(x_n - x_n')$$

Direct Product Spaces (cont.)

Or, if we want to check the commutators we postulated in Section 6.5, we have

$$\left\langle \prod_{p=1}^{N} x_p \right| \left[X_j^{\prod_{m=1}^{N}(m)}, P_j^{\prod_{n=1}^{N}(n)} \right] \left| \prod_{q=1}^{N} x_q' \right\rangle$$

$$= \langle x_j \left| \left[X_j^{(j)}, P_j^{(j)} \right] \right| x_j' \rangle \prod_{n \neq j}^{N} \langle x_n | I^{(n)} | x_n' \rangle = i\hbar \delta(x_j - x_j') \prod_{n \neq j}^{N} \delta(x_n - x_n')$$

$$= i\hbar \prod_{n=1}^{N} \delta(x_n - x_n') \tag{11.2}$$

and, for $j \neq k$,

$$\left\langle \prod_{p=1}^{N} x_p \right| \left[X_j^{\prod_{m=1}^{N}(m)}, P_k^{\prod_{n=1}^{N}(n)} \right] \left| \prod_{q=1}^{N} x_q' \right\rangle$$

$$= \left[\langle x_j | X_j^{(j)} | x_j' \rangle \langle x_k | P_k^{(k)} | x_k' \rangle - \langle x_k | P_k^{(k)} | x_k' \rangle \langle x_j | X_j^{(j)} | x_j' \rangle \right] \prod_{n \neq j,k}^{N} \langle x_n | I^{(n)} | x_n' \rangle = 0$$

because the inner products $\langle x_j | X_j^{(j)} | x_j' \rangle$ and $\langle x_k | P_k^{(k)} | x_k' \rangle$ commute because they are just numbers. One could prove $\left[X_j^{\prod_{m=1}^{N}(m)}, X_k^{\prod_{n=1}^{N}(n)} \right] = 0$ and $\left[P_j^{\prod_{m=1}^{N}(m)}, P_k^{\prod_{n=1}^{N}(n)} \right] = 0$ by a similar technique.

Time Evolution

For a direct product space $\mathbb{V} = \prod_{n=1}^{N} \mathbb{V}_n$, there is a direct product space Hamiltonian operator H that gives the time evolution via the Schrödinger Equation:

$$i\,\hbar\,\frac{d}{dt}\,|\psi\rangle^{\prod_{n=1}^{N}(n)} = H\,|\psi\rangle^{\prod_{n=1}^{N}(n)}$$

Note that it is in general not possible to write H in a simpler form; in particular, it is in general not possible to factor H into single-dof operators, $H = H^{\prod_{m=1}^{N}(m)}$, as we shall see later.

The Hamiltonian and Time-Evolution (cont.)

As usual, H will have eigenstates that satisfy

$$H|\psi_E\rangle\prod_{n=1}^{N}(n) = E\,|\psi_E\rangle\prod_{n=1}^{N}(n)$$

Because H is the Hamiltonian for the whole system, E is the energy for the whole system. For these energy eigenstates, the Schrödinger Equation is

$$i\,\hbar\,\frac{d}{dt}\,|\psi_E\rangle\prod_{n=1}^{N}(n) = E\,|\psi_E\rangle\prod_{n=1}^{N}(n)$$

and the time evolving solution is

$$|\psi_E(t)\rangle\prod_{n=1}^{N}(n) = e^{-\frac{i}{\hbar}\,E\,t}|\psi_E\rangle\prod_{n=1}^{N}(n)$$

Separable vs. Non-Separable Hamiltonians

A separable Hamiltonian is one that can be written in the form

$$H^{\prod_{n=1}^{N}(n)} = \sum_{n=1}^{N} H^{(n)} \otimes \prod_{m \neq n} I^{(m)} \tag{11.3}$$

The eigenvalue-eigenvector equation simplifies to a set of single-dof Schrödinger equations in this case. We can see this as follows. First, write down the eigenvalue-eigenvector equation:

$$E \, |\psi_E\rangle^{\prod_{q=1}^{N}(q)} = H^{\prod_{n=1}^{N}(n)} |\psi_E\rangle^{\prod_{q=1}^{N}(q)} = \sum_{n=1}^{N} \left(H^{(n)} \otimes \prod_{m \neq n} I^{(m)} \right) |\psi_E\rangle^{\prod_{q=1}^{N}(q)}$$

The Hamiltonian and Time-Evolution (cont.)

Now, expand the direct product space eigenstate in a basis consisting of direct products of eigenstates of the individual single-dof Hamiltonians. We are assured that the latter are valid bases for the individual factor spaces $\mathbb{V}_n$, so we know from our earlier discussion that the direct products of these basis elements form a basis for the direct product space. That is, assume there exist states $\{|\psi_{E_m^{(n)}}\rangle^{(n)}\}$ in $\mathbb{V}_n$ that satisfy

$$H^{(n)}\,|\psi_{E_m^{(n)}}\rangle^{(n)} = E_m^{(n)}\,|\psi_{E_m^{(n)}}\rangle^{(n)}$$

where m tells us which eigenstate of $H^{(n)}$ we are referring to. The $^{(n)}$ superscript in $E_m^{(n)}$ is necessary because E_m alone does not tell us which Hamiltonian's eigenvalues we are referring to (*i.e.*, which of the $H^{(n)}$ is E_m the eigenvalue for?). Then our basis for the direct product space consists of direct product states of the form

$$\prod_{n=1}^{N} |\psi_{E_{m_n}^{(n)}}\rangle^{(n)}$$

where m_n is the index telling one that the basis state uses the m_n'th eigenstate of $H^{(n)}$ (yes, unfortunately, the $_n$ subscript in m_n is required to be clear as we will see below).

The Hamiltonian and Time-Evolution (cont.)

For the sake of notational brevity, we will write the above as

$$\prod_{n=1}^{N} |\psi_{E_{m_n}}\rangle$$

because the $_n$ subscript on m_n tells us the same thing as the $^{(n)}$ superscripts do: that E_{m_n} is an eigenvalue and that $|\psi_{E_{m_n}}\rangle$ is an eigenvector of $H^{(n)}$.

Consider the action of the Hamiltonian on a state of this type:

$$\begin{aligned}
H^{\prod_{n=1}^{N}(n)} |\psi_E\rangle^{\prod_{q=1}^{N}(q)} &= \sum_{n=1}^{N} \left(H^{(n)} \otimes \prod_{q\neq n} I^{(q)} \right) \prod_{k=1}^{N} |\psi_{E_{m_k}}\rangle \\
&= \sum_{n=1}^{N} \left[\left(H^{(n)} |\psi_{E_{m_n}}\rangle \right) \otimes \prod_{q\neq n} \left(I^{(q)} |\psi_{E_{m_q}}\rangle \right) \right] \\
&= \sum_{n=1}^{N} E_{m_n} \left(\prod_{q=1}^{N} |\psi_{E_{m_q}}\rangle \right) = \left(\sum_{n=1}^{N} E_{m_n} \right) \left(\prod_{q=1}^{N} |\psi_{E_{m_q}}\rangle \right) \\
&\equiv E_{m_1\cdots m_N} |\psi_E\rangle^{\prod_{q=1}^{N}(q)} \quad \text{with} \quad E_{m_1\cdots m_N} = E_{m_1} + E_{m_2} + \cdots + E_{m_N}
\end{aligned}$$

The Hamiltonian and Time-Evolution (cont.)

The time evolution of the eigenstates of the direct product Hamiltonian are therefore

$$|\psi_{E_{m_1\cdots m_N}}(t)\rangle\prod_{n=1}^{N}(n) = e^{-\frac{i}{\hbar}E_{m_1\cdots m_N}t}|\psi_{E_{m_1\cdots m_N}}\rangle\prod_{n=1}^{N}(n)$$
$$= e^{-\frac{i}{\hbar}E_{m_1\cdots m_N}t}\left(\prod_{n=1}^{N}|\psi_{E_{m_n}}\rangle\right) = \left(\prod_{n=1}^{N}e^{-\frac{i}{\hbar}E_{m_n}t}|\psi_{E_{m_n}}\rangle\right)$$

One just has the direct product of the time-evolving individual single-dof eigenstates. Note that this implies that the energies of the individual particles (or individual degrees of freedom) are independently conserved.

One can already see the opportunity for energy degeneracies in the direct product space when there is no such degeneracy in the single-dof states because the total energy is the sum of the individual energies; if the individual degrees of freedom are similar (e.g., all SHOs), then one will have many combinations of single-dof states that yield the same energy. This energy degeneracy is critical to statistical mechanics — the idea that many microstates (sets of individual particle energies) correspond to the same macrostate (total energy).

Now, the above is of course a fairly boring case because there are no interactions between the different degrees of freedom. In particular, if H includes a potential term, the above says that the potential is not just factorizable, but it is a sum of independent terms in the different degrees of freedom.

Lecture 31:
Multiparticle Systems
Separable vs. Non-Separable Hamiltonians ct'd
Position-Space Wavefunctions

Date Revised: 2009/01/09
Date Given: 2009/01/09

Example 11.1: Separable and non-separable combinations of two single-dof potentials

For a particle in two spatial dimensions, the potential $V(x_1, x_2)$ in a separable Hamiltonian must be of the form $V(x_1) + V(x_2)$; even $V(x_1, x_2) = V_1(x_1)V_2(x_2)$ is not separable. So, for example, $V(x_1, x_2) = x_1^2 + x_2^4$ is separable, but $V(x_1, x_2) = x_1^2\, x_2^4$ is not, and certainly $V(x_1, x_2) = \sqrt{x_1^2 + x_2^4}$ is not.

We can see this fact by writing the potentials in the formally correct way as operators in the direct product space and seeing whether the operator can be written in separable form or not. The matrix elements of the potential operator in the direct product space can always be written as

$$ {}^{(1)\otimes(2)}\langle x_1, x_2 \,|\, V^{(1)\otimes(2)} | x_1', x_2' \,\rangle^{(1)\otimes(2)} = V(x_1, x_2)\, \delta(x_1 - x_1')\, \delta(x_2 - x_2') $$

where $V(x_1, x_2)$ are the simple functions (not operators) defined above. It is natural to define the single-particle potential operators

$$ {}^{(1)}\langle x_1 \,|\, V^{(1)} | x_1' \,\rangle^{(1)} = V_1(x_1)\delta(x_1 - x_1') = x_1^2 \delta(x_1 - x_1') $$
$$ {}^{(2)}\langle x_2 \,|\, V^{(2)} | x_2' \,\rangle^{(2)} = V_2(x_2)\delta(x_2 - x_2') = x_2^4 \delta(x_2 - x_2') $$

The Hamiltonian and Time-Evolution (cont.)

With those definitions, the first potential can be written

$$(x_1^2 + x_2^2)\,\delta(x_1 - x_1')\,\delta(x_2 - x_2') = {}^{(1)\otimes(2)}\langle x_1, x_2 \,|\, \left[V^{(1)} \otimes I^{(2)} + I^{(1)} \otimes V^{(2)}\right] |\, x_1', x_2' \,\rangle^{(1)\otimes(2)}$$

The potential operator is in this case separable into two terms that are each the identity operator in all of the factor spaces except one. The second potential can be written

$$(x_1^2 x_2^2)\,\delta(x_1 - x_1')\,\delta(x_2 - x_2') = {}^{(1)\otimes(2)}\langle x_1, x_2 \,|\, \left[V^{(1)} \otimes V^{(2)}\right] |\, x_1', x_2' \,\rangle^{(1)\otimes(2)}$$

In this case, the potential operator is factorizable into a direct product of two single-particle operators, but that is not a separable operator.

The third potential cannot be decomposed in either of these simpler fashions.

Example 11.2: Dependence of separability on choice of generalized coordinates

The 2D simple oscillator, which can also be thought of as two independent 1D simple harmonic oscillators, is separable:

$$V(x_1, x_2) = \frac{1}{2} k x_1^2 + \frac{1}{2} k x_2^2$$

or, more formally,

$$V^{(1)\otimes(2)} = \left(\frac{1}{2} k X_1^2\right) \otimes I + I \otimes \left(\frac{1}{2} k X_2^2\right)$$

where we have left out the $^{(1)}$ and $^{(2)}$ subscripts when possible.

The Hamiltonian and Time-Evolution (cont.)

There is another separable way to write the above potential. Classically, we know the above potential can be separated in polar coordinates:

$$V(r,\phi) = \frac{1}{2}\, k\, r^2$$

If we choose our factor Hilbert spaces to describe the polar coordinates r and ϕ instead of x_1 and x_2, then, instead of factor spaces $\mathbb{V}^{(1)}$ and $\mathbb{V}^{(2)}$ describing the behavior in x_1 and x_2, we have factor spaces $\mathbb{V}^{(r)}$ and $\mathbb{V}^{(\phi)}$ describing the behavior in r and ϕ. The potential is not only separable, but it also consists of only one term:

$$V^{(r)\otimes(\phi)} = \left(\frac{1}{2}\, k\, R^2\right) \otimes I^{(\phi)}$$

The particle is free in the ϕ coordinate and looks like a 1D SHO restricted to $r > 0$ and with a boundary condition at $r = 0$.

The Hamiltonian and Time-Evolution (cont.)

The potential

$$V(x_1, x_2) = \sqrt{\frac{1}{2}\, k\, x_1^2 + \frac{1}{2}\, k\, x_2^2}$$

is clearly non-separable (and non-factorizable) in x_1 and x_2, but it can be rewritten in polar coordinates as

$$V(r, \phi) = \left(\frac{1}{2}\, k\right)^{1/2} r$$

is again separable in r and ϕ. Here, as with the Coulomb potential between two particles, separability depends on the choice of the classical generalized coordinates.

Example 11.3: Separability by change of variables

Some potentials are non-separable when written down in their obvious form as an operator in a direct product space whose factor spaces are single-particle Cartesian degrees of freedom. However, by changing the the degrees of freedom we use to describe the system, and hence the factor Hilbert spaces, we can sometimes obtain some amount of separability. A classic example of this type is the Coulomb interaction between two particles. The Hamiltonian in the direct product space of the single-particle Hilbert spaces is

$$H = \frac{P_{1,x}^2 + P_{1,y}^2 + P_{1,z}^2}{2\,m_1} + \frac{P_{2,x}^2 + P_{2,y}^2 + P_{2,z}^2}{2\,m_2} - \frac{e^2}{\sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2 + (Z_1 - Z_2)^2}}$$

where every operator is implied to have a $^{(1)\otimes(2)}$ superscript and where the second term is a shorthand for implying what the potential term's matrix elements are in the position basis. While the kinetic terms are trivially separable into single-particle Hamiltonians, the potential term is definitely not.

The Hamiltonian and Time-Evolution (cont.)

However, in this particular case, the problem can be rewritten into new degrees of freedom that are, to some extent, separable. Define

$$X_{CM} = \frac{m_1 X_1 + m_2 X_2}{M} \qquad P_{CM,x} = P_{1,x} + P_{2,x}$$

$$X_{12} = X_1 - X_2 \qquad P_{12,x} = \frac{m_2}{M} P_{1,x} + \frac{m_1}{M} P_{2,x} \qquad M = m_1 + m_2$$

and so on for the y and z degrees of freedom. Then the Hamiltonian becomes

$$H = \frac{P_{CM,x}^2 + P_{CM,y}^2 + P_{CM,z}^2}{2\,M} + \frac{P_{12,x}^2 + P_{12,y}^2 + P_{12,z}^2}{2\,\mu} - \frac{e^2}{\sqrt{X_{12}^2 + Y_{12}^2 + Z_{12}^2}}$$

$$\equiv H_{CM,x} + H_{CM,y} + H_{CM,z} + H_{12}$$

$$\mu = \frac{m_1\,m_2}{M}$$

The Hamiltonian and Time-Evolution (cont.)

In doing the above, we have implicitly chosen new factor spaces: rather than considering the states of particle 1 and particle 2 separately, we consider the state of the center of mass coordinate and the difference coordinate. This should be somewhat disconcerting, as it is the first time we have considered the idea of the states in a Hilbert space representing anything besides the state of a single particle. But the redefinition is completely reasonable; the first Hilbert space is the state describing the system's center of mass (where it is, how fast it moves, etc.) and the second is the one describing the different coordinate (how far apart the particles are, how quickly this distance changes, etc.). In quantum mechanics, we use Hilbert space states to represent the behavior of a system in its classical generalized coordinates, which may or may not correspond to single-particle degrees of freedom.

With the above, we see that, if, instead of considering our full Hilbert space to be the direct product of individual states of the two particles in three-dimensional space, and thus of six one-dimensional degrees of freedom, we consider it to be the direct product of a one free three-dimensional particle (*i.e.*, three one-dimensional free degrees of freedom) and one three-dimensional particle subject to a Coulomb interaction, then the problem is partially separable. We will see later that H_{12} can be separated further by again re-choosing the classical degrees of freedom to be spherical rather than cartesian coordinates.

The above separability proof holds for any two-particle Hamiltonian in which the potential depends only on the vector separation of the two particles. In the above case, the potential in fact only depends on distance between the two particles.

Position-Space Wavefunctions

Since we frequently work with particle states in the position basis — *i.e.*, the position-space wavefunction — let us be explicit about how to project direct product states onto the position basis of the direct product space.

First, what is the position basis of the direct product space? It is the direct product of the individual position bases:

$$|x_1, \cdots, x_N\rangle^{\prod_{n=1}^{N}(n)} = \prod_{n=1}^{N} |x_n\rangle^{(n)}$$

This is a complete basis for the direct product space because the factors are complete bases for the factor spaces.

Position-Space Wavefunction (cont.)

Now, let's write the state $|\psi\rangle^{\prod_{n=1}^{N}(n)}$ in terms of the position basis. That is, there are some set of coefficients for expanding $|\psi\rangle^{\prod_{n=1}^{N}(n)}$ in terms of the $\{|x_1, \cdots, x_N\rangle^{\prod_{n=1}^{N}(n)}\}$; there have to be, since the latter are a complete basis. Let's label those coefficients $\psi(x_1, \cdots, x_N)$:

$$
\begin{aligned}
&|\psi\rangle^{\prod_{n=1}^{N}(n)} \\
&\quad = \int_{-\infty}^{\infty} dx_1' \int_{-\infty}^{\infty} dx_2' \cdots \int_{-\infty}^{\infty} dx_N' \, \psi(x_1', x_2', \cdots, x_N') \, |x_1', x_2', \cdots, x_N'\rangle^{\prod_{n=1}^{N}(n)} \\
&\qquad = \int_{-\infty}^{\infty} dx_1' \int_{-\infty}^{\infty} dx_2' \cdots \int_{-\infty}^{\infty} dx_N' \, \psi(x_1', x_2', \cdots, x_N') \prod_{n=1}^{N} |x_n'\rangle^{(n)}
\end{aligned}
$$

Position-Space Wavefunction (cont.)

Then, the position space wavefunction is obtained trivially:

$$
\begin{aligned}
&\textstyle\prod_{m=1}^{N}{}^{(m)}\langle x_1, \cdots, x_N \,|\psi\rangle \textstyle\prod_{n=1}^{N}{}^{(n)} \\
&= \left(\prod_{m=1}^{N}{}^{(m)}\langle x_m\,|\right) \int_{-\infty}^{\infty} dx_1' \int_{-\infty}^{\infty} dx_2' \cdots \int_{-\infty}^{\infty} dx_N'\, \psi(x_1', x_2', \cdots, x_N') \prod_{n=1}^{N} |x_n'\rangle^{(n)} \\
&= \int_{-\infty}^{\infty} dx_1' \int_{-\infty}^{\infty} dx_2' \cdots \int_{-\infty}^{\infty} dx_N'\, \psi(x_1', x_2', \cdots, x_N') \prod_{n=1}^{N} {}^{(n)}\langle x_n \,|x_n'\rangle^{(n)} \\
&= \int_{-\infty}^{\infty} dx_1' \int_{-\infty}^{\infty} dx_2' \cdots \int_{-\infty}^{\infty} dx_N'\, \psi(x_1', x_2', \cdots, x_N') \prod_{n=1}^{N} \delta(x_n - x_n') \\
&= \psi(x_1, \cdots, x_N)
\end{aligned}
$$

Let's consider the implications of the above wavefunction for position measurement. Since $|x_1, \cdots, x_N\rangle \prod_{m=1}^{N}{}^{(m)}$ projects onto the state in which degree of freedom (dof) 1 is at x_1, dof 2 is at x_2, etc., the squared modulus of the position-space wavefunction $|\psi(x_1, \cdots, x_N)|^2$ gives the probability of measuring dof 1 to be at x_1, dof 2 to be at x_2, and so on. The position space wavefunction thus gives the probability of a particular position configuration of the ensemble of degrees of freedom.

Position-Space Wavefunction (cont.)

In some cases, the direct product space state may be factorizable into a product of single-dof states:

$$|\psi\rangle^{\prod_{n=1}^{N}(n)} = \prod_{n=1}^{N} |\psi_n\rangle^{(n)}$$

In such cases, the position-space wavefunction can be seen to be factorizable:

$$\psi(x_1, \cdots, x_N) = \left(\prod_{m=1}^{N} {}^{(m)}\langle x_m|\right) |\psi\rangle^{\prod_{n=1}^{N}(n)} = \left(\prod_{m=1}^{N} {}^{(m)}\langle x_m|\right)\left(\prod_{n=1}^{N} |\psi_n\rangle^{(n)}\right)$$
$$= \prod_{n=1}^{N} {}^{(n)}\langle x_n|\psi_n\rangle^{(n)} = \prod_{n=1}^{N} \psi_n(x_n)$$

where we have used the definition of the single-dof position-space wavefunction, $\psi_n(x_n) = {}^{(n)}\langle x_n|\psi_n\rangle^{(n)}$. This form is one you may be familiar with from previous courses. Here, we have that the direct product space position wavefunction is, literally, the arithmetic product of the individual dof position-space wavefunctions.

Position-Space Wavefunction (cont.)

This simple outcome — arithmetic product instead of direct product — arises because the position-space wavefunction is not the Hilbert space state itself, but gives the expansion coefficients in terms of the position basis. The coefficients of such expansions are always numbers (regardless of the basis in terms of which one expands), hence we expect to end up with arithmetic multiplication of numbers instead of direct product.

The above concept of course generalizes to any basis: momentum-basis, energy-basis, etc. In all cases, the wavefunction for that basis, which is of course just the coefficients of the expansion of the state in that basis, gives the probability of that particular configuration for the ensemble of the degrees of freedom.

One may even end up with "hybrid" wavefunctions. As we will see later, we describe particles with intrinsic spin as living in the direct product space of a standard particle Hilbert space (three dof) and a Hilbert space that describes spin (it turns out to be a vector space of dimension $2j+1$ where j is the spin). One might choose to use the position basis for the spatial dof; for the spin dof, one typically chooses a basis of S_z, spin z-projection, eigenstates (see Examples 3.3, 3.7, 3.14). The expansion coefficients in the resulting direct product basis can be thought of as a "position-space/spin-space" wavefunction.

Lecture 32:
Indistinguishable Particles

Date Revised: 2009/01/12
Date Given: 2009/01/12

The Basic Problem

We have assumed that the degrees of freedom under study are distinguishable: we can label them and identify them. For example, we call one "1" and the other "2", and no matter where we observe them in whatever space they live in, it is clear which is which. This is trivial if the degrees of freedom are different — *e.g.*, two different spatial dimensions for a single particle, two different kinds of degrees of freedom for an object (position and rigid body orientation), etc. It is nontrivial if the degrees of freedom are the same — *e.g.*, two particles in the same single spatial dimension.

In particular, fundamental particles are in general indistinguishable — one electron is indistinguishable from another because an electron has no internal degrees of freedom that can be used to label it (we neglect particle spin for now.)

So what? Our construction of the position or momentum operator for a particular particle in the direct product space is now called into question. Recall, we defined

$$X_1^{(1)\otimes(2)} \equiv X_1^{(1)} \otimes I^{(2)}$$

But, if we cannot distinguish particle 1 from particle 2, how do we know whether our "X" measurement is a measurement of particle 1 or particle 2?

We must revisit our construction of direct product space states and operators — *i.e.*, Postulates 1, 2, and 3 — when we consider systems of indistinguishable particles.

States for Indistinguishable Particles

Let's first revisit Postulate 1. We need the states to live in a Hilbert space. But nothing says that we must take the entire direct product Hilbert space to be our Hilbert space.

Let's first elucidate better how the indistinguishability of the particles relates the two factor spaces. The factor spaces $\mathbb{V}_1$ and $\mathbb{V}_2$ must be identical if the particles are indistinguishable. Note that we are not saying they are the same space; but rather that they are identical to each other in every way, in the same way that the x and y axes in two dimensions are identical but are separate spaces.

One implication is that for any state $|v\rangle^{(1)}$ in $\mathbb{V}_1$, there is a matching state $|v\rangle^{(2)}$ in $\mathbb{V}_2$ and vice versa. This is not necessarily true for distinguishable particles or distinguishabel degrees of freedom.

Operators for Indistinguishable Particles

Given that $\mathbb{V}_1$ and $\mathbb{V}_2$ are identical, we recognize that for any operator $\Omega^{(1)}$ acting in $\mathbb{V}_1$, there is a matching operator $\Omega^{(2)}$ acting in $\mathbb{V}_2$ that acts in an identical manner on $\mathbb{V}_2$ as $\Omega^{(1)}$ acts on $\mathbb{V}_1$. What do we mean by "identical"? An operator is a rule assigning to any input ket an output ket. We've already stated that for any $|v\rangle^{(1)}$ in $\mathbb{V}_1$, there is a corresponding $|v\rangle^{(2)}$ in $\mathbb{V}_2$. $\Omega^{(1)}|v\rangle^{(1)}$ is another ket $|w\rangle^{(1)}$ in $\mathbb{V}_1$. By saying that $\mathbb{V}_1$ and $\mathbb{V}_2$ are identical, we are assured there is a state $|w\rangle^{(2)}$ that matches up with $|w\rangle^{(1)}$. But the operator $\Omega^{(2)}$ acting on $|v\rangle^{(2)}$ generates some output state $\Omega^{(2)}|v\rangle^{(2)}$ in $\mathbb{V}_2$. When we say that the action of $\Omega^{(2)}$ in $\mathbb{V}_2$ is identical to that of $\Omega^{(1)}$ in $\mathbb{V}_1$, we are saying:

$$\text{Given} \quad |v\rangle^{(1)} \leftrightarrow |v\rangle^{(2)}, \quad |w\rangle^{(1)} = \Omega^{(1)}|v\rangle^{(1)}, \quad \Omega^{(1)} \leftrightarrow \Omega^{(2)}, \quad \text{and} \quad |w\rangle^{(1)} \leftrightarrow |w\rangle^{(2)},$$
$$\text{then} \quad |w\rangle^{(2)} = \Omega^{(2)}|v\rangle^{(2)}$$

($\leftrightarrow$ indicates "correspondence"). One could exchange the last condition and the inference, but the point is the same: there is a matching of the action of the operators.

Back to States for Indistinguishable Particles

The above discussion of operators implies that, given two eigenvalues $\widehat{\omega}$ and $\widetilde{\omega}$, $\widehat{\omega} \neq \widetilde{\omega}$ and corresponding eigenstates $|\widehat{\omega}\rangle$ and $|\widetilde{\omega}\rangle$ for a single-particle operator Ω, the state $|\widehat{\omega}, \widetilde{\omega}\rangle^{(1)\otimes(2)}$ is an eigenstate for both $\Omega^{(1)}$ and $\Omega^{(2)}$ with eigenvalues $\widehat{\omega}$ and $\widetilde{\omega}$ for particles 1 and 2, respectively. But, if the particles are indistinguishable, there is no experimental way to create a state that has particle 1 is in state $|\widehat{\omega}\rangle$ and particle 2 in state $|\widetilde{\omega}\rangle$. We must consider linear combinations of the form

$$\alpha\,|\widehat{\omega}, \widetilde{\omega}\rangle^{(1)\otimes(2)} + \beta\,|\widetilde{\omega}, \widehat{\omega}\rangle^{(1)\otimes(2)}$$

Furthermore, we must require that the state be unchanged under exchange of the two particles' eigenvalues because, if it were not, then the particles would be distinguishable. So, we require

$$\alpha\,|\widetilde{\omega}, \widehat{\omega}\rangle^{(1)\otimes(2)} + \beta\,|\widehat{\omega}, \widetilde{\omega}\rangle^{(1)\otimes(2)} = \gamma\left[\alpha\,|\widehat{\omega}, \widetilde{\omega}\rangle^{(1)\otimes(2)} + \beta\,|\widetilde{\omega}, \widehat{\omega}\rangle^{(1)\otimes(2)}\right]$$

Indistinguishable Particles (cont.)

We thus obtain the conditions

$$\alpha = \gamma\,\beta \qquad \beta = \gamma\,\alpha \qquad \Longrightarrow \qquad \gamma^2 = 1 \qquad \alpha^2 = \beta^2$$

Note that we have the square, not the modulus squared, of each variable, which implies that the solutions are

$$\gamma = 1 \quad \beta = \alpha \qquad \text{and} \qquad \gamma = -1 \quad \beta = -\alpha$$

So the allowed states are (up to normalization)

$$|\widehat{\omega}, \widetilde{\omega}\,\rangle_{+}^{(1)\otimes(2)} = |\widehat{\omega}, \widetilde{\omega}\,\rangle_{S}^{(1)\otimes(2)} = |\widehat{\omega}, \widetilde{\omega}\,\rangle^{(1)\otimes(2)} + |\widetilde{\omega}, \widehat{\omega}\,\rangle^{(1)\otimes(2)}$$
$$|\widehat{\omega}, \widetilde{\omega}\,\rangle_{-}^{(1)\otimes(2)} = |\widehat{\omega}, \widetilde{\omega}\,\rangle_{A}^{(1)\otimes(2)} = |\widehat{\omega}, \widetilde{\omega}\,\rangle^{(1)\otimes(2)} - |\widetilde{\omega}, \widehat{\omega}\,\rangle^{(1)\otimes(2)}$$

These states are symmetric and antisymmetric under exchange of the two particles.

Since the original "distinguishable particle" states $|\widehat{\omega}, \widetilde{\omega}\,\rangle^{(1)\otimes(2)}$ form a basis for the space (because the eigenstates of any Hermitian operator are a basis), the ensemble of both symmetric and antisymmetric states also form a basis for the space.

Indistinguishable Particles (cont.)

However, we may not allow both kinds of states in our Hilbert space since. If both exist, we can construct the "distinguishable states" by linear combinations:

$$\frac{1}{2}\left(|\widehat{\omega},\widetilde{\omega}\,\rangle_{+}^{(1)\otimes(2)} + |\widehat{\omega},\widetilde{\omega}\,\rangle_{-}^{(1)\otimes(2)}\right) = |\widehat{\omega},\widetilde{\omega}\,\rangle^{(1)\otimes(2)}$$

This is to be expected, as the ensemble of symmetric and antisymmetric states form a basis for the whole product space. So, what we must do is restrict to either the symmetric or antisymmetric part of the basis. By restricting the basis, we restrict the entire space, of course; we *restrict the physical Hilbert space* to contain either symmetric or antisymmetric states. We shall see later that these two kinds of particles are called bosons and fermions and that the above requirement has significant physical consequences.

Back to Operators for Indistinguishable Particles

The following discussion is not present in Shankar.

Now, we have the problem that our allowed states are no longer eigenstates of our operators. For example:

$$\begin{aligned}
\Omega_1^{(1)\otimes(2)} |\widehat{\omega}, \widetilde{\omega}\rangle_{\pm}^{(1)\otimes(2)} &= \left(\Omega_1^{(1)} \otimes I^{(2)}\right) |\widehat{\omega}, \widetilde{\omega}\rangle_{\pm}^{(1)\otimes(2)} \\
&= \left(\Omega_1^{(1)} \otimes I^{(2)}\right) \left(|\widehat{\omega}, \widetilde{\omega}\rangle^{(1)\otimes(2)} \pm |\widetilde{\omega}, \widehat{\omega}\rangle^{(1)\otimes(2)}\right) \\
&= \widehat{\omega} |\widehat{\omega}, \widetilde{\omega}\rangle^{(1)\otimes(2)} \pm \widetilde{\omega} |\widetilde{\omega}, \widehat{\omega}\rangle^{(1)\otimes(2)} \\
&= \frac{\widehat{\omega}}{2} \left(|\widehat{\omega}, \widetilde{\omega}\rangle_{+}^{(1)\otimes(2)} + |\widehat{\omega}, \widetilde{\omega}\rangle_{-}^{(1)\otimes(2)}\right) \\
&\quad \pm \frac{\widetilde{\omega}}{2} \left(|\widehat{\omega}, \widetilde{\omega}\rangle_{+}^{(1)\otimes(2)} - |\widehat{\omega}, \widetilde{\omega}\rangle_{-}^{(1)\otimes(2)}\right) \\
&= \frac{\widetilde{\omega} \pm \widehat{\omega}}{2} |\widehat{\omega}, \widetilde{\omega}\rangle_{+}^{(1)\otimes(2)} + \frac{\widetilde{\omega} \mp \widehat{\omega}}{2} |\widehat{\omega}, \widetilde{\omega}\rangle_{-}^{(1)\otimes(2)}
\end{aligned}$$

It makes sense that there is a problem because the above operator distinguishes the particles by picking one to operate on. We need new operators.

Indistinguishable Particles (cont.)

How do we construct operators that respect the indistinguishable nature of the particles? Recall that, in the case of distinguishable particles, the operator on the product space to measure some observable ω for particle 1 would be $\Omega_1^{(1)\otimes(2)} = \Omega^{(1)} \otimes I^{(2)}$ and for particle 2 would be $\Omega_2^{(1)\otimes(2)} = I^{(1)} \otimes \Omega^{(2)}$. So, to construct the indistinguishable particle analogue, we simply average and difference the two:

$$\overline{\Omega}^{(1)\otimes(2)} = \frac{1}{2}\left(\Omega_1^{(1)\otimes(2)} + \Omega_2^{(1)\otimes(2)}\right) = \frac{1}{2}\left(\Omega^{(1)} \otimes I^{(2)} + I^{(1)} \otimes \Omega^{(2)}\right)$$

$$\left[\delta\Omega^{(1)\otimes(2)}\right]^2 = \left[\Omega_1^{(1)\otimes(2)} - \Omega_2^{(1)\otimes(2)}\right]^2 = \left[\Omega^{(1)} \otimes I^{(2)} - I^{(1)} \otimes \Omega^{(2)}\right]^2$$

That is, we construct the two linear combinations of $\Omega_1^{(1)\otimes(2)}$ and $\Omega_2^{(1)\otimes(2)}$ that are symmetric under exchange of the two particles. We will see below why we need the absolute value for the difference operator. Note that we mathematically allow the "distinguishable particle operators" $\Omega_1^{(1)\otimes(2)}$ and $\Omega_2^{(1)\otimes(2)}$, but we do not allow them physically: any indistinguishable particle observable must be made of a linear combination of the above type.

The action of the above operators on our allowed states is then

$$\begin{aligned}
\overline{\Omega}^{(1)\otimes(2)}|\widehat{\omega},\widetilde{\omega}\rangle_{\pm}^{(1)\otimes(2)} &= \frac{1}{2}\left(\Omega^{(1)}\otimes I^{(2)} + I^{(1)}\otimes\Omega^{(2)}\right)\left(|\widehat{\omega},\widetilde{\omega}\rangle^{(1)\otimes(2)} \pm |\widetilde{\omega},\widehat{\omega}\rangle^{(1)\otimes(2)}\right) \\
&= \frac{\widehat{\omega}+\widetilde{\omega}}{2}|\widehat{\omega},\widetilde{\omega}\rangle^{(1)\otimes(2)} \pm \frac{\widetilde{\omega}+\widehat{\omega}}{2}|\widehat{\omega},\widetilde{\omega}\rangle^{(1)\otimes(2)} \\
&= \frac{\widehat{\omega}+\widetilde{\omega}}{2}|\widehat{\omega},\widetilde{\omega}\rangle_{\pm}^{(1)\otimes(2)} \\
\left[\delta\Omega^{(1)\otimes(2)}\right]^2|\widehat{\omega},\widetilde{\omega}\rangle_{\pm}^{(1)\otimes(2)} &= \left[\Omega^{(1)}\otimes I^{(2)} - I^{(1)}\otimes\Omega^{(2)}\right]^2\left(|\widehat{\omega},\widetilde{\omega}\rangle^{(1)\otimes(2)} \pm |\widetilde{\omega},\widehat{\omega}\rangle^{(1)\otimes(2)}\right) \\
&= (\widehat{\omega}-\widetilde{\omega})^2|\widehat{\omega},\widetilde{\omega}\rangle^{(1)\otimes(2)} \pm (\widetilde{\omega}-\widehat{\omega})^2|\widehat{\omega},\widetilde{\omega}\rangle^{(1)\otimes(2)} \\
&= (\widehat{\omega}-\widetilde{\omega})^2|\widehat{\omega},\widetilde{\omega}\rangle_{\pm}^{(1)\otimes(2)}
\end{aligned}$$

In both cases, the expression did not vanish for the $(-)$ case because the two states being differenced are different states. We see why we needed the absolute value for the difference operator: had we note taken the absolute value, the sign of the eigenvalue would have distinguished $|\widehat{\omega},\widetilde{\omega}\rangle_{\pm}^{(1)\otimes(2)}$ from $|\widetilde{\omega},\widehat{\omega}\rangle_{\pm}^{(1)\otimes(2)}$.

Bosons and Fermions

We have explained about how we must restrict the physical Hilbert space to either the symmetric or antisymmetric piece for indistinguishable particles. Since any sum of symmetric states is also a symmetric state, and similarly for antisymmetric states, we see that the each of these restricted spaces is a subspace, and hence we may write the direct product as a direct sum:

$$\mathbb{V}_{(1)\otimes(2)} = \mathbb{V}_+ \oplus \mathbb{V}_- = \mathbb{V}_S \oplus \mathbb{V}_A$$

Recall that the above does *not* imply that $\mathbb{V}_{(1)\otimes(2)}$ is the union of the two subspaces: the direct sum spaces is bigger than either of the subspaces $\mathbb{V}_+$ and $\mathbb{V}_-$ because there are linear combinations of symmetric and antisymmetric states – such as the $|\widehat{\omega}, \widetilde{\omega}\,\rangle^{(1)\otimes(2)}$ example considered earlier — that do not belong to either subspace. The result is that the restriction to $\mathbb{V}_+$ or $\mathbb{V}_-$ significantly reduces the number of states available. For Hilbert spaces that have countably or uncountably infinite dimension, this reduction means little, but the idea is the same.

An analogy would be think of $\mathbb{V}_{(1)\otimes(2)}$ as the real plane, $\mathbb{V}_+$ as all the vectors along the line $y = x$, and $\mathbb{V}_-$ as all the vectors along the line $y = -x$. The two spaces $\mathbb{V}_+$ and $\mathbb{V}_-$ are subspaces and their linear combinations give the entire space $\mathbb{V}_{(1)\otimes(2)}$, but their union would just be the points satisfying $y = \pm x$.

Indistinguishable Particles (cont.)

One might be tempted to think that $\mathbb{V}_+$ and $\mathbb{V}_-$ are in some sense identical, that for every state in $\mathbb{V}_+$ there is a state in $\mathbb{V}_-$ and vice versa. This turns out not to be true: consider the symmetric and antisymmetric versions of the "distinguishable particle state" $|\omega,\omega\rangle^{(1)\otimes(2)}$. It is clear that

$$|\omega,\omega\rangle_+^{(1)\otimes(2)} = |\omega,\omega\rangle^{(1)\otimes(2)}$$
$$|\omega,\omega\rangle_-^{(1)\otimes(2)} = 0$$

(where 0 means the null vector here). That is, there is a subset of states in $\mathbb{V}_+$ for which there is no partner in $\mathbb{V}_-$.

We note that the above statement is the Pauli Exclusion Principle: two indistinguishable particles cannot be in the same state if they are fermions because the particle-exchange-antisymmetrized state is the null state. There is no such prohibition on bosons since such states are automatically symmetric under particle exchange.

Expansions in Terms of Symmetric and Antisymmetric Bases

In constructing states in terms of $\mathbb{V}_+$ or $\mathbb{V}_-$ basis elements, we need to be careful to avoid double-counting. For distinguishable particles, our generic 2-particle state is of the form (see Equation 11.1)

$$|u\rangle^{(1)\otimes(2)} = \sum_{j,k} u_{jk} |v_j, w_k\rangle^{(1)\otimes(2)}$$

where there is no restriction on j and k. If we specialize to a system with two particles that are identical but distinguishable, the above becomes

$$|u\rangle^{(1)\otimes(2)} = \sum_{j,k} u_{jk} |\omega_j, \omega_k\rangle^{(1)\otimes(2)}$$

where we have labeled the single particle states by ω_j and ω_k instead of v_j and w_k because the two factor spaces are now identical (*i.e.*, we should use v_k instead of w_k) and for consistency with our early notation (use ω_j, ω_k instead of v_j, v_k). There is no restriction on j and k yet.

Indistinguishable Particles (cont.)

Let's now rewrite in terms of the symmetric and antisymmetric bases, still treating the particles as distinguishable:

$$
\begin{aligned}
|u\rangle^{(1)\otimes(2)} &= \sum_j u_{jj}\,|\omega_j,\omega_j\rangle^{(1)\otimes(2)} \\
&\quad + \sum_{j>k}\left[u_{jk}\,|\omega_j,\omega_k\rangle^{(1)\otimes(2)} + u_{kj}\,|\omega_k,\omega_j\rangle^{(1)\otimes(2)}\right] \\
&= \sum_j u_{jj}|\omega_j,\omega_j\rangle^{(1)\otimes(2)} \\
&\quad + \sum_{j>k}\Bigg[\frac{u_{jk}}{\sqrt{2}}\left(|\omega_j,\omega_k\rangle_+^{(1)\otimes(2)} + |\omega_j,\omega_k\rangle_-^{(1)\otimes(2)}\right) \\
&\qquad\quad + \frac{u_{kj}}{\sqrt{2}}\left(|\omega_j,\omega_k\rangle_+^{(1)\otimes(2)} - |\omega_j,\omega_k\rangle_-^{(1)\otimes(2)}\right)\Bigg]
\end{aligned}
$$

Now, define

$$
u_{jk,+} = \frac{u_{jk}+u_{kj}}{\sqrt{2}} \qquad |\omega_j,\omega_k\rangle_+^{(1)\otimes(2)} = \frac{1}{\sqrt{2}}\left(|\omega_j,\omega_k\rangle^{(1)\otimes(2)} + |\omega_k,\omega_j\rangle^{(1)\otimes(2)}\right)
$$

$$
u_{jk,-} = \frac{u_{jk}-u_{kj}}{\sqrt{2}} \qquad |\omega_j,\omega_k\rangle_-^{(1)\otimes(2)} = \frac{1}{\sqrt{2}}\left(|\omega_j,\omega_k\rangle^{(1)\otimes(2)} - |\omega_k,\omega_j\rangle^{(1)\otimes(2)}\right)
$$

Indistinguishable Particles (cont.)

Therefore, we have

$$
\begin{aligned}
&= \sum_j u_{jj} \, |\omega_j, \omega_j\rangle^{(1)\otimes(2)} \\
&\quad + \sum_{j>k} \left(\frac{u_{jk} + u_{kj}}{\sqrt{2}} \, |\omega_j, \omega_k\rangle_+^{(1)\otimes(2)} + \frac{u_{jk} - u_{kj}}{\sqrt{2}} \, |\omega_j, \omega_k\rangle_-^{(1)\otimes(2)} \right) \\
&\equiv \sum_j u_{jj} \, |\omega_j, \omega_j\rangle^{(1)\otimes(2)} + \sum_{j>k} \left(u_{jk,+} \, |\omega_j, \omega_k\rangle_+^{(1)\otimes(2)} + u_{jk,-} \, |\omega_j, \omega_k\rangle_-^{(1)\otimes(2)} \right)
\end{aligned}
$$

We have pulled out the $j = k$ part of the sum because there is no antisymmetric contribution to it and because the state does not get a $1/\sqrt{2}$ in front. We have properly normalized the $+$ and $-$ states assuming the "distinguishable particle" states are normalized. We explicitly see the decomposition $\mathbb{V}_{(1)\otimes(2)} = \mathbb{V}_+ \oplus \mathbb{V}_-$. Also, now instead of coefficients u_{jk} with $j > k$ and $j < k$ allowed, we consider only the coefficients $u_{jk,+}$ and $u_{jk,-}$ with $j > k$ allowed. Same number of free parameters, but labeled in a different way.

Indistinguishable Particles (cont.)

Now, if we consider identical, indistinguishable particles, we need to restrict to states that are either purely symmetric or antisymmetric. In the above, this is simple: for symmetric states, we set all $u_{jk,-} = 0$, and for antisymmetric states, we set all $u_{jj} = 0$ and all $u_{jk,+} = 0$. The resulting two kinds of states are

$$|u\rangle_+^{(1)\otimes(2)} = \sum_j u_{jj} |\omega_j, \omega_j\rangle^{(1)\otimes(2)} + \sum_{j>k} u_{jk,+} |\omega_j, \omega_k\rangle_+^{(1)\otimes(2)}$$
$$|u\rangle_-^{(1)\otimes(2)} = \sum_{j>k} u_{jk,-} |\omega_j, \omega_k\rangle_-^{(1)\otimes(2)}$$

If the $|\omega_j, \omega_k\rangle^{(1)\otimes(2)}$ have unity norm, then $|\omega_j, \omega_j\rangle^{(1)\otimes(2)}$ and $|\omega_j, \omega_k\rangle_\pm^{(1)\otimes(2)}$ have unity norm, so the normalization condition on the states is

$$\sum_j |u_{jj}|^2 + \sum_{j>k} |u_{jk,+}|^2 = 1 \qquad \text{or} \qquad \sum_{j>k} |u_{jk,-}|^2 = 1 \tag{11.4}$$

So, the algorithm for normalizing the states is clear.

Lecture 33:
Multiparticle Systems:
Indistinguishable Particles continued

Date Revised: 2009/01/16
Date Given: 2009/01/14

Expectation Values

Are expectation values for observables normalized correctly? Shankar goes through some rigamarole with factors of 1/2 to get this all straight. But if we use the "indistinguishable particle" operators we constructed earlier, then everything is properly normalized. We can check this explicitly:

$$\begin{aligned}\langle \overline{\Omega}^{(1)\otimes(2)} \rangle_+ &= {}^{(1)\otimes(2)}_{+}\langle u \,|\, \overline{\Omega}^{(1)\otimes(2)} | u \,\rangle^{(1)\otimes(2)}_{+} \\ &= \sum_j |u_{jj}|^2 \, {}^{(1)\otimes(2)}_{+}\langle \omega_j, \omega_j \,|\, \overline{\Omega}^{(1)\otimes(2)} | \omega_j, \omega_j \,\rangle^{(1)\otimes(2)}_{+} \\ &\quad + \sum_{j>k} |u_{jk,+}|^2 \, {}^{(1)\otimes(2)}_{+}\langle \omega_j, \omega_k \,|\, \overline{\Omega}^{(1)\otimes(2)} | \omega_j, \omega_k \,\rangle^{(1)\otimes(2)}_{+} \end{aligned} \tag{11.5}$$

Recall that $|\omega_j, \omega_k\,\rangle^{(1)\otimes(2)}_{\pm}$ is an eigenvector of $\overline{\Omega}$ with eigenvalue $(\omega_j + \omega_k)/2$. So the sum becomes easy:

$$\langle \overline{\Omega}^{(1)\otimes(2)} \rangle_+ = \sum_j |u_{jj}|^2 \omega_j + \sum_{j>k} |u_{jk,+}|^2 \left(\frac{\omega_j + \omega_k}{2} \right)$$

Indistinguishable Particles (cont.)

For an antisymmetric state, we have

$$\langle \overline{\Omega}^{(1)\otimes(2)} \rangle_- = \sum_{j>k} |u_{jk,-}|^2 \; {}^{(1)\otimes(2)}_{-}\langle \omega_j, \omega_k \, | \, \overline{\Omega}^{(1)\otimes(2)} | \omega_j, \omega_k \, \rangle^{(1)\otimes(2)}_{-}$$

$$= \sum_{j>k} |u_{jk,-}|^2 \left(\frac{\omega_j + \omega_k}{2} \right) \tag{11.6}$$

For the difference operator, we have

$$\left\langle \left[\delta\Omega^{(1)\otimes(2)} \right]^2 \right\rangle_+ = \sum_j |u_{jj}|^2(0) + \sum_{j>k} |u_{jk,+}|^2(\omega_j - \omega_k)^2 = \sum_{j>k} |u_{jk,+}|^2(\omega_j - \omega_k)^2$$

$$\left\langle \left[\delta\Omega^{(1)\otimes(2)} \right]^2 \right\rangle_- = \sum_{j>k} |u_{jk,-}|^2(\omega_j - \omega_k)^2$$

Indistinguishable Particles (cont.)

It is sometimes more convenient to let both j and k run over their full range and throw in a factor of $1/2$ for the double counting because the full sums, or, for a continuous eigenvalue, the full integrals, are easier to evaluate than the partial sums. That is, we may rewrite Equation 11.5 as

$$\begin{aligned}
\langle \overline{\Omega}^{(1)\otimes(2)} \rangle_+ &= \sum_j |u_{jj}|^2 \; {}^{(1)\otimes(2)}_{+}\langle \omega_j, \omega_j \,|\, \overline{\Omega}^{(1)\otimes(2)} | \omega_j, \omega_j \,\rangle^{(1)\otimes(2)}_{+} \\
&\qquad + \sum_{j>k} |u_{jk,+}|^2 \; {}^{(1)\otimes(2)}_{+}\langle \omega_j, \omega_k \,|\, \overline{\Omega}^{(1)\otimes(2)} | \omega_j, \omega_k \,\rangle^{(1)\otimes(2)}_{+} \\
&= \frac{1}{2} \sum_j |u_{jj,+}|^2 \; {}^{(1)\otimes(2)}_{+}\langle \omega_j, \omega_j \,|\, \overline{\Omega}^{(1)\otimes(2)} | \omega_j, \omega_j \,\rangle^{(1)\otimes(2)}_{+} \\
&\qquad + \frac{1}{2} \sum_{j\neq k} |u_{jk,+}|^2 \; {}^{(1)\otimes(2)}_{+}\langle \omega_j, \omega_k \,|\, \overline{\Omega}^{(1)\otimes(2)} | \omega_j, \omega_k \,\rangle^{(1)\otimes(2)}_{+} \\
&= \frac{1}{2} \sum_{j,k} |u_{jk,+}|^2 \; {}^{(1)\otimes(2)}_{+}\langle \omega_j, \omega_k \,|\, \overline{\Omega}^{(1)\otimes(2)} | \omega_j, \omega_k \,\rangle^{(1)\otimes(2)}_{+}
\end{aligned}$$

where we have defined $u_{jj,+}$ by extending the formula for $u_{jk,+}$ without modification, $u_{jj,+} = \frac{u_{jj}+u_{jj}}{\sqrt{2}} = \sqrt{2}\, u_{jj}$.

Indistinguishable Particles (cont.)

and 11.6 as

$$\begin{aligned}
\langle \overline{\Omega}^{(1)\otimes(2)} \rangle_- &= \sum_{j>k} |u_{jk,-}|^2 \, {}^{(1)\otimes(2)}_{-}\langle \omega_j, \omega_k \,|\, \overline{\Omega}^{(1)\otimes(2)} | \omega_j, \omega_k \,\rangle^{(1)\otimes(2)}_{-} \\
&= \frac{1}{2} \sum_{j\neq k} |u_{jk,-}|^2 \, {}^{(1)\otimes(2)}_{-}\langle \omega_j, \omega_k \,|\, \overline{\Omega}^{(1)\otimes(2)} | \omega_j, \omega_k \,\rangle^{(1)\otimes(2)}_{-} \\
&= \frac{1}{2} \sum_{j,k} |u_{jk,-}|^2 \, {}^{(1)\otimes(2)}_{+}\langle \omega_j, \omega_k \,|\, \overline{\Omega}^{(1)\otimes(2)} | \omega_j, \omega_k \,\rangle^{(1)\otimes(2)}_{-}
\end{aligned}$$

where $u_{jj,-} = 0$ follows from the definition of $u_{jk,-}$, $u_{jj,-} = \frac{u_{jj} - u_{jj}}{\sqrt{2}} = 0$.

If the eigenvalue ω is continuous and we write the above as integrals (with the original single-particle states $|\omega\rangle$ appropriately normalized given that ω is now continuous), we have

$$\begin{aligned}
\langle \overline{\Omega}^{(1)\otimes(2)} \rangle_+ &= \frac{1}{2} \int d\widehat{\omega} \int d\widetilde{\omega} \, |u_+(\widehat{\omega}, \widetilde{\omega})|^2 \, {}^{(1)\otimes(2)}_{+}\langle \widehat{\omega}, \widetilde{\omega} \,|\, \overline{\Omega}^{(1)\otimes(2)} | \widehat{\omega}, \widetilde{\omega} \,\rangle^{(1)\otimes(2)}_{+} \\
\langle \overline{\Omega}^{(1)\otimes(2)} \rangle_- &= \frac{1}{2} \int d\widehat{\omega} \int d\widetilde{\omega} \, |u_-(\widehat{\omega}, \widetilde{\omega})|^2 \, {}^{(1)\otimes(2)}_{-}\langle \widehat{\omega}, \widetilde{\omega} \,|\, \overline{\Omega}^{(1)\otimes(2)} | \widehat{\omega}, \widetilde{\omega} \,\rangle^{(1)\otimes(2)}_{-}
\end{aligned}$$

which are Shankar's equations 10.3.15 and 10.3.25 with $\widehat{\omega} = x_1$ and $\widetilde{\omega} = x_2$.

Indistinguishable Particles (cont.)

Probabilities

If we do a measurement of $\overline{\Omega}$ followed by a measurement $(\delta\Omega)^2$, we will obtain values $\frac{\omega_j+\omega_k}{2}$ and $(\omega_j-\omega_k)^2$ ($\frac{\widehat{\omega}+\widetilde{\omega}}{2}$ and $(\widehat{\omega}-\widetilde{\omega})^2$ for the continuous case). (Note that $\overline{\Omega}$ and $(\delta\Omega)^2$ commute and are hence simultaneously diagonalizable.) From these, we may calculate the two numbers ω_j and ω_k (or $\widehat{\omega}$ and $\widetilde{\omega}$), though, we don't know which is which; that is, we obtain

$$\omega_j \quad \text{or} \quad \widehat{\omega} = \langle\overline{\Omega}\rangle \pm \sqrt{\left\langle(\delta\Omega)^2\right\rangle} \qquad \omega_k \quad \text{or} \quad \widetilde{\omega} = \langle\overline{\Omega}\rangle \mp \sqrt{\left\langle(\delta\Omega)^2\right\rangle}$$

and we don't know which sign pair to choose because of the indistinguishability of the particles. What is the probability of obtaining this measurement, which we shall designate by $P_\pm(\widehat{\omega},\widetilde{\omega})$, where the $\pm$ reminds us that the two values are unordered because of the above indeterminacy? The above expressions for the state normalizations (Equation 11.4) and for the expectation values $\langle\overline{\Omega}\rangle_\pm$ and $\left\langle(\delta\Omega)^2\right\rangle_\pm$ make this clear: the probability is

$$P_\pm(\omega_j,\omega_k) = |u_{jk,\pm}|^2 \qquad \text{or} \qquad P_\pm(\widehat{\omega},\widetilde{\omega}) = |u_\pm(\widehat{\omega},\widetilde{\omega})|^2$$

We are assured that the probability, summed or integrated over all possible outcomes, is normalized to unity by Equation 11.4.

Indistinguishable Particles (cont.)

Example 11.4: Two free particles of the same mass with a particle-in-a-box attractive potential between them. Consider the cases of the particles being distinguishable and indistinguishable.

Consider two free particles on the real line with position coordinates x_1 and x_2 and suppose they experience the following attractive potential:

$$V(x_1, x_2) = \begin{cases} 0 & |x_1 - x_2| < L/2 \\ \infty & |x_1 - x_2| \geq L/2 \end{cases}$$

What are the eigenstates and energy eigenvalues of the system?

Let's begin by writing down the full Hamiltonian in the two-particle direct product space, where $\mathbb{V}_1$ and $\mathbb{V}_2$ are the individual Hilbert spaces for the two particles separately.

$$H^{(1)\otimes(2)} = \frac{1}{2\,m}\left[P_1^{(1)\otimes(2)}\right]^2 + \frac{1}{2\,m}\left[P_2^{(1)\otimes(2)}\right]^2 V^{(1)\otimes(2)}$$

$$\text{where} \quad P_1^{(1)\otimes(2)} = P^{(1)} \otimes I^{(2)} \qquad P_2^{(1)\otimes(2)} = I^{(1)} \otimes P^{(2)}$$

$$\text{and} \quad {}^{(1)\otimes(2)}\langle x_1, x_2 \,|\, V^{(1)\otimes(2)} | x_1', x_2' \,\rangle^{(1)\otimes(2)} = V(x_1, x_2)\,\delta(x_1 - x_1')\,\delta(x_2 - x_2')$$

The kinetic energy part of the Hamiltonian is that of two free particles, but the potential energy is not, rendering the Hamiltonian unseparable.

Indistinguishable Particles (cont.)

The clear solution is to change to CM and difference coordinates. One can do this in one of two ways. The first is to rewrite the classical Hamiltonian in terms of CM and difference coordinates and then take those difference coordinates over to operators in QM that satisfy the canonical commutation relations. The other thing way to do it, which is more straightforward here, is to define new QM operators in terms of the old ones. That is

$$X_{CM}^{(1)\otimes(2)} = \frac{1}{2}\left(X_1^{(1)\otimes(2)} + X_2^{(1)\otimes(2)}\right) \qquad X_{12}^{(1)\otimes(2)} = X_1^{(1)\otimes(2)} - X_2^{(1)\otimes(2)}$$

$$P_{CM}^{(1)\otimes(2)} = \frac{1}{2}\left(P_1^{(1)\otimes(2)} + P_2^{(1)\otimes(2)}\right) \qquad P_{12}^{(1)\otimes(2)} = P_1^{(1)\otimes(2)} - P_2^{(1)\otimes(2)}$$

One can check that the new operators satisfy the usual commutation relations (dropping the now unnecessary $^{(1)\otimes(2)}$ superscript):

$$[X_{CM}, P_{CM}] = i\,\hbar \qquad [X, P] = i\,\hbar$$

$$[X_{CM}, P_{12}] = 0 = [X_{12}, P_{CM}] \qquad [X_{12}, X_{CM}] = 0 = [P_{12}, P_{CM}]$$

Indistinguishable Particles (cont.)

The new operators have their own eigenbases, which we can express in terms of the original bases as follows:

$$|x_{CM}, x_{12}\rangle^{(CM)\otimes(12)} = \left|x_{CM} + \frac{x_{12}}{2}, x_{CM} - \frac{x_{12}}{2}\right\rangle^{(1)\otimes(2)}$$

$$|p_{CM}, p_{12}\rangle^{(CM)\otimes(12)} = \left|p_{CM} + \frac{p_{12}}{2}, p_{CM} - \frac{p_{12}}{2}\right\rangle^{(1)\otimes(2)}$$

We have introduced the $^{(CM)\otimes(12)}$ superscripting not because the kets on the left side live in a different Hilbert space, but rather to make it clear whether a ket $|x_a, x_b\rangle$ refers to an eigenstate of X_{CM} and X_{12} or an eigenstate of X_1 and X_2. It is not an unreasonable notation, however, because one could have begun with single-dof Hilbert spaces $\mathbb{V}_{CM}$ and $\mathbb{V}_{12}$ if one had changed to CM and relative coordinates in the classical variables first. The Hamiltonian becomes

$$H = \frac{1}{2\,M}\,P_{CM}^2 + \frac{1}{2\,\mu}\,P^2 + V$$

The first two terms are clearly separable in the new labeling of the basis, but that was not the problem; it was the V term.

Indistinguishable Particles (cont.)

Let's evaluate the matrix elements of V in our new labeling of the position basis by using the correspondence to the original labeling:

$$
\begin{aligned}
&{}^{(CM)\otimes(12)}\langle x_{CM}, x_{12}|V|x'_{CM}, x'_{12}\rangle^{(CM)\otimes(12)} \\
&= {}^{(1)\otimes(2)}\left\langle x_1 = x_{CM} + \frac{x_{12}}{2}, x_2 = x_{CM} - \frac{x_{12}}{2}\right| V \left|x_1' = x'_{CM} + \frac{x'_{12}}{2}, x_2' = x'_{CM} - \frac{x'_{12}}{2}\right\rangle \\
&\quad = \delta\left(\left[x_{CM} + \frac{x_{12}}{2}\right] - \left[x'_{CM} + \frac{x'_{12}}{2}\right]\right)\delta\left(\left[x_{CM} - \frac{x_{12}}{2}\right] - \left[x'_{CM} - \frac{x'_{12}}{2}\right]\right) \\
&\qquad\qquad \times \begin{cases} 0 & |x_{12}| < L/2 \\ \infty & |x_{12}| \geq L/2 \end{cases}
\end{aligned}
$$

where we have used the same evaluation for the matrix element of V as we did when we initially wrote down the Hamiltonian, but now we have replaced x_1, x_1', x_2, and x_2' with x_{CM}, x'_{CM}, x_{12}, and x'_{12}.

Indistinguishable Particles (cont.)

Let's evaluate the product of delta functions. The expression is nonzero only when the arguments of both delta functions vanish, which occurs when

$$x_{CM} - x'_{CM} = \frac{1}{2}\left(x'_{12} - x_{12}\right) \quad \text{and} \quad x_{CM} - x'_{CM} = \frac{1}{2}\left(x_{12} - x'_{12}\right)$$

Therefore, we may replace $x_{CM} - x'_{CM}$ in the second delta function with the value given by the first equation above (or, alternately, replace it in the first delta function with the value in the second equation). This gives us for the delta function product

$$\begin{aligned}\delta\left(\left[x_{CM} + \frac{x_{12}}{2}\right] - \left[x'_{CM} + \frac{x'_{12}}{2}\right]\right)\, \delta\left(\frac{1}{2}\left[x'_{12} - x_{12}\right] + \frac{1}{2}\left[x'_{12} - x_{12}\right]\right) \\ = \delta\left(\left[x_{CM} - x'_{CM}\right] + \frac{1}{2}\left[x_{12} - x'_{12}\right]\right)\, \delta(x'_{12} - x_{12})\end{aligned}$$

Using the same kind of manipulation, we can set $x_{12} - x'_{12}$ to zero in the first delta function, yielding

$$\delta(x_{CM} - x'_{CM})\,\delta(x_{12} - x'_{12})$$

Note that we reversed the sign of the argument in the second delta function, which is allowed because a delta function is an even functions of its argument.

Indistinguishable Particles (cont.)

Therefore, we may rewrite the matrix element for the potential operator as

$$ {}^{(CM)\otimes(12)}\langle x_{CM}, x_{12}\,|V|x'_{CM}, x'_{12}\,\rangle^{(CM)\otimes(12)} = $$
$$ \delta(x_{CM}-x'_{CM})\,\delta(x_{12}-x'_{12}) \times \begin{cases} 0 & |x_{12}| < L/2 \\ \infty & |x_{12}| \geq L/2 \end{cases} $$

In the end, we have relabeled our basis so the space is of the form

$$ \mathbb{V} = \mathbb{V}_{CM} \otimes \mathbb{V}_{12} $$

and the Hamiltonian breaks into two pieces,

$$ H = H^{(CM)} \otimes I^{(12)} + I^{(CM)} \otimes H^{(12)} $$
$$ \text{with} \quad H^{(CM)} \otimes I^{(12)} = P^2_{CM} \quad \text{and} \quad I^{(CM)} \otimes H^{(12)} = P^2_{12} + V $$

where the P^2_{CM} operator is the identity in the $\mathbb{V}_{12}$ factor space, the P^2_{12} operator is the identity in the $\mathbb{V}_{CM}$ factor space, and the V operator has explicitly been shown to be the identity in the $\mathbb{V}_{CM}$ factor space.

Indistinguishable Particles (cont.)

The Hamiltonian is thus separable, giving us a free particle in x_{CM} and a particle in a box in x_{12}. We know how to solve these separate single-particle problems, so we may conclude that the eigenvalues of the Hamiltonian are

$$E = E_{CM\pm} + E_{12} \quad \text{with} \quad E_{CM\pm} = \text{any value} \geq 0$$
$$\text{and} \quad E_{12} = \frac{\hbar^2 \pi^2 n^2}{2 \mu L^2} \quad n > 0 \quad \text{any integer}$$

and the projections of the eigenstates onto the position basis (the position-basis wavefunctions) are

$$\psi_{(E_{CM\pm},E_{12}),(x_{CM},x_{12})}(x_{CM}, x_{12}) = {}^{(CM)\otimes(12)}\langle x_{CM}, x_{12} | E_{CM\pm}, E_{12}\rangle$$
$$= \frac{1}{\sqrt{2\pi\hbar}} \sqrt{\frac{2}{L}}\, e^{\pm \frac{i}{\hbar}\sqrt{2 M E_{CM\pm}}\, x_{CM}} \begin{cases} \cos \frac{n \pi x_{12}}{L} & n > 0, \text{ odd} \\ \sin \frac{n \pi x_{12}}{L} & n > 0, \text{ even} \end{cases}$$

The $\pm$ in $E_{CM\pm}$ refers to whether the solution is a right-going or left-going plane wave.

Indistinguishable Particles (cont.)

How do we modify the above for indistinguishable particles? One way to do this would be to replace X_1 and X_2 with indistinguishable particle operators and repeat the process of recasting the Hamiltonian in terms of CM and difference coordinate operators. This would be quite straightforward, as $\overline{P} = P_{CM}$, $(\delta P)^2 = P_{12}^2$, and V would become a function of $\sqrt{(\delta X)^2}$. We would solve the problem in the symmetrized space $\mathbb{V}_+$ or antisymmetrized space $\mathbb{V}_-$, meaning that, instead of calculating the matrix elements of the Hamiltonian in the basis

$$|x_{CM}, x_{12}\rangle^{(CM)\otimes(12)} = \left|x_1 = x_{CM} + \frac{x_{12}}{2}, x_2 = x_{CM} - \frac{x_{12}}{2}\right\rangle^{(1)\otimes(2)}$$

we would do the calculation in the symmetrized or antisymmetrized basis

$$\frac{1}{\sqrt{2}}\left(|x_{CM}, x_{12}\rangle^{(CM)\otimes(12)} \pm |x_{CM}, -x_{12}\rangle^{(CM)\otimes(12)}\right)$$
$$= \frac{1}{\sqrt{2}}\left(\left|x_1 = x_{CM} + \frac{x_{12}}{2}, x_2 = x_{CM} - \frac{x_{12}}{2}\right\rangle^{(1)\otimes(2)}\right.$$
$$\left. \pm \left|x_2 = x_{CM} - \frac{x_{12}}{2}, x_1 = x_{CM} + \frac{x_{12}}{2}\right\rangle^{(1)\otimes(2)}\right)$$

Indistinguishable Particles (cont.)

We could write our vector space as the direct sum

$$\mathbb{V} = \mathbb{V}_+ \oplus \mathbb{V}_-$$

but let's use the factorizability of the vector space into *CM* and difference coordinates as we did before. The *CM* coordinate does not care about the sign of x_{12}, so which subspace we pick does not affect the choice of which *CM* coordinate states are allowed. Thus, we may conclude that the space can be written as

$$\mathbb{V} = \mathbb{V}_{CM} \otimes (\mathbb{V}_{12,+} \oplus \mathbb{V}_{12,-}) = (\mathbb{V}_{CM} \otimes \mathbb{V}_{12,+}) \oplus (\mathbb{V}_{CM} \otimes \mathbb{V}_{12,-})$$

and that we need to pick one of the two subspaces to work in. Our basis elements factorize:

$$\frac{1}{\sqrt{2}} \left(|x_{CM}, x_{12}\rangle^{(CM)\otimes(12)} \pm |x_{CM}, -x_{12}\rangle^{(CM)\otimes(12)} \right)$$
$$= |x_{CM}\,\rangle^{(CM)} \otimes \frac{1}{\sqrt{2}} \left(|x_{12}\,\rangle^{(12)} \pm |-x_{12}\,\rangle^{(12)} \right)$$

Indistinguishable Particles (cont.)

We could go through and re-solve the single particle in a box in the $\mathbb{V}_{12,+}$ or $\mathbb{V}_{12,-}$ subspace — that is, project the eigenvalue-eigenvector equation onto symmetrized or antisymmetrized versions of the $|x_{12}\rangle$ basis for $\mathbb{V}_{12}$, which would restrict us to $\mathbb{V}_{12,+}$ or $\mathbb{V}_{12,-}$ as required. But it is easier to simply note that our existing solutions already are restricted to these subspaces: the cos solutions live in the $\mathbb{V}_{12,+}$ subspace and the sin solutions live in the $\mathbb{V}_{12,-}$ subspace. Thus, we need only pick one or the other type of solution based on whether our particles are bosons (n odd) or fermions (n even).

We could have seen this more quickly, though perhaps less rigorously, by simplying seeing from the start that the n odd states are allowed for bosons and the n even states are allowed for fermions because these pick up the appropriate sign under $x_{12} \longrightarrow -x_{12}$, which is equivalent to exchanging the particles. x_{CM} is unaffected by particle exchange, so the part of the position-basis wavefunction that depends on x_{CM} is unchanged by exchange.

Lecture 34:
Multiparticle Systems:
Indistinguishable Particles continued

Date Revised: 2009/01/18
Date Given: 2009/01/16

Is it a Boson or a Fermion?

How do we experimentally determine whether to use the symmetric or antisymmetric states for a particle? It turns out it is easy because of the propensity that fermions have to avoid the same state. We put two particles in a box in single-particle eigenstates $|\psi_a\rangle$ and $|\psi_b\rangle$. Then we calculate what the probability of finding them at the position pair $(\widehat{x}, \widetilde{x})$ is, depending on their statistics. Using the above formula for the probability of obtaining a particular position pair outcome, we have

$$P_{\pm}(\widehat{x}, \widetilde{x}) = |u_{\pm}(\widehat{x}, \widetilde{x})|^2$$

What is $u_{\pm}(\widehat{x}, \widetilde{x})$? It is the projection of the system state onto the symmetric or antisymmetric position eigenstate $|\widehat{x}, \widetilde{x}\rangle_{\pm}$. That is,

$$u_{\pm}(\widehat{x}, \widetilde{x}) = \quad {}_{\pm}\langle \widehat{x}, \widetilde{x} | \psi_a, \psi_b \rangle_{\pm}$$

Indistinguishable Particles (cont.)

Let's expand this out in terms of distinguishable particle eigenstates for calculational convenience (remembering that such states are not physically allowed!):

$$\begin{aligned}
{}_{\pm}\langle \widehat{x}, \widetilde{x} \,|\psi_a, \psi_b\,\rangle_{\pm} &= \frac{1}{\sqrt{2}}\left(\langle \widehat{x}, \widetilde{x}\,| \pm \langle \widetilde{x}, \widehat{x}\,|\right) \frac{1}{\sqrt{2}}\left(|\psi_a, \psi_b\,\rangle \pm |\psi_b, \psi_a\,\rangle\right) \\
&= \frac{1}{2}\Big[\psi_{a,x}(\widehat{x})\,\psi_{b,x}(\widetilde{x}) + \psi_{b,x}(\widetilde{x})\psi_{a,x}(\widehat{x}) \\
&\qquad \pm \psi_{b,x}(\widehat{x})\,\psi_{a,x}(\widetilde{x}) \pm \psi_{a,x}(\widetilde{x})\,\psi_{b,x}(\widehat{x})\Big] \\
&= \psi_{a,x}(\widehat{x})\,\psi_{b,x}(\widetilde{x}) \pm \psi_{b,x}(\widehat{x})\,\psi_{a,x}(\widetilde{x})
\end{aligned}$$

where we made the last step because the functions $\psi_{a,x}$ and $\psi_{b,x}$ are just numbers now and therefore commute.

Now, we calculate the probability:

$$\begin{aligned}
P_{\pm}(\widehat{x}, \widetilde{x}) = |u_{\pm}(\widehat{x}, \widetilde{x})|^2 &= \left[\psi_{a,x}(\widehat{x})\,\psi_{b,x}(\widetilde{x}) \pm \psi_{b,x}(\widehat{x})\,\psi_{a,x}(\widetilde{x})\right]^* \qquad (11.7) \\
&\qquad \times \left[\psi_{a,x}(\widehat{x})\,\psi_{b,x}(\widetilde{x}) \pm \psi_{b,x}(\widehat{x})\,\psi_{a,x}(\widetilde{x})\right] \\
&= |\psi_{a,x}(\widehat{x})|^2|\psi_{b,x}(\widetilde{x})|^2 + |\psi_{b,x}(\widehat{x})|^2|\psi_{a,x}(\widetilde{x})|^2 \qquad (11.8) \\
&\qquad \pm \psi^*_{a,x}(\widehat{x})\,\psi_{b,x}(\widehat{x})\,\psi^*_{b,x}(\widetilde{x})\,\psi_{a,x}(\widetilde{x}) \\
&\qquad \pm \psi^*_{b,x}(\widehat{x})\,\psi_{a,x}(\widehat{x})\,\psi^*_{a,x}(\widetilde{x})\,\psi_{b,x}(\widetilde{x})
\end{aligned}$$

Indistinguishable Particles (cont.)

That is, we obtain entirely different probability distributions depending on whether the particles are bosons or fermions. And, moreover, had the particles been distinguishable, but we just didn't know how to distinguish them, so that the particles were in an *incoherent* superposition of (a, b) and (b, a), the cross terms with the $\pm$ would not have appeared at all. So we can even distinguish distinguishable particles from indistinguishable bosons and fermions, even if we didn't know they were distinguishable! An example would be two electrons in opposite spin states, but supposing we didn't know about spin.

It is interesting to calculate $P_\pm(x_0, x_0)$, the probability of finding the two particles at the same position x_0:

$$\begin{aligned} P_\pm(x_0, x_0) &= 2\left(|\psi_{a,x}(x_0)|^2|\psi_{b,x}(x_0)|^2 \pm |\psi_{a,x}(x_0)|^2|\psi_{b,x}(x_0)|^2\right) \\ &= \begin{cases} 4\,|\psi_{a,x}(x_0)|^2|\psi_{b,x}(x_0)|^2 & + \\ 0 & - \end{cases} \end{aligned}$$

and we would have obtained $P(x_0, x_0) = 2\,|\psi_{a,x}(x_0)|^2|\psi_{b,x}(x_0)|^2$ if the particles were distinguishable. The above result is inevitable from Postulate 3: the position measurements would put the two particles into position eigenstates, but there is no antisymmetric state with both particles at the same position, so the probability for obtaining that measurement outcome had better vanish!

Indistinguishable Particles (cont.)

So we have a clear experimental means for deciding the indistinguishability characeristics for new particles. See also the example in Shankar about the K and $\overline{K}$ mesons.

We note in passing that the above implies that a plot of $P_-(\widehat{x}, \widetilde{x})$ as a function of $\widehat{x}$ and $\widetilde{x}$ would be a surface with a node along the line $\widehat{x} = \widetilde{x}$. $P(\widehat{x}, \widetilde{x})$, and a plot of $P_+(\widehat{x}, \widetilde{x})$ would be larger along this line than a plot of $P(\widehat{x}, \widetilde{x})$ (which would of course be larger than $P_-(\widehat{x}, \widetilde{x})$ along this line.)

Indistinguishable Particles (cont.)

Symmetrization and Antisymmetrization for Systems of N Particles

Suppose one has N indistinguishable particles. What should the properties of the overall state be under permutation of the particles?

The key to answering this question is to realize that any permutation of N objects can be obtained by successive pairwise exchanges. One can prove this inductively.

Suppose you have N objects, n of which are out of their ordered positions. Pick the first one that is out of its normal place i; call it i and suppose it is sitting in position j. Clearly, position i must be occupied by another object k whose normal position is $k \neq i$ because i belongs in position i. So, exchange these two particles, so that i is now in position i and k is in position j. One either has now $n-1$ (if $k \neq j$) or $n-2$ (if $k = j$) objects out of their normal positions. So, a permutation with n out of N objects out of place can be obtained by pairwise exchange from a permutation with either $n-1$ or $n-2$ objects out of place.

Now, prove it for $n = 2$ and $n = 3$ explicitly so the inductive proof carries through. For $n = 2$, just exchange the two out of place objects — done. For $n = 3$, label the three objects a, b, and c. Exchange a and b. Exactly one of them must now be in its correct position: first, if both were still of out place, then they both have only one other choice for their correct positions, the position that c is in, and that must be wrong because two objects cannot have the same correct position; if both were now in their correct positions, then object c must also now be in its correct position, which is wrong by assumption that $n = 3$. So, if $n = 3$, a single pair exchange brings us to $n = 2$. We know how to send $n = 2$ to $n = 0$ by pair exchange. QED.

Indistinguishable Particles (cont.)

Once one sees that any permutation of N particles can be obtained by pair exchange, one understands that one only has to apply the pair exchange rule multiple times. For bosons, the state should be an equally weighted sum over all permutations of the particles with no sign flips. For fermions, the state is an equally weighted sum over all permutations also, but with the "odd" permutations — those that require an odd number of pair exchanges to obtain — picking up -1's and the "even" permutations retaining $+1$ coefficients. Assuming the single-particle states are normalized, the whole thing requires a $\sqrt{N!}$ normalization factor. If we designate the jth of $N!$ possible permutations as P_j, where $m = P_j(k)$ indicates that particle k in the unpermuted state should be replaced by particle m in the jth permuted state, and the permutations are defined so that P_j requires an odd number of exchanges when j is odd and an even number when j is even, then we have

$$|\psi_{n_1}, \cdots, \psi_{n_N}\rangle_{\pm} = \frac{1}{\sqrt{N!}} \sum_{j=1}^{N!} (\pm 1)^j |\psi_{n_1}\rangle^{(P_j(1))} \otimes \cdots \otimes |\psi_{n_N}\rangle^{(P_j(N))}$$

$$= \frac{1}{\sqrt{N!}} \sum_{j=1}^{N!} (\pm 1)^j \prod_{k=1}^{N} |\psi_{n_k}\rangle^{(P_j(k))}$$

where, in order to provide compact notation, the states are allowed to go out of particle order; the $^{(P_j(k))}$ indicates which particle the state n_k is assigned to in the jth permutation.

Indistinguishable Particles (cont.)

The above notation works only if all the states are distinct (if $n_k \neq n_m$ for all k and m) or if the particles are fermions. If the particles are fermions, the above formula will vanish if there are duplicate states, which is fine. For bosons, if there are duplicate states, there will be terms in the sum that repeat. For example, if two of the states are identical, $n_k = n_m$ for some k, m pair, then, for every permutation j, there will be a matching permutation that differs only by exchanging the two particles $P_j(k)$ and $P_j(m)$ to which the states $n_k = n_m$ have been assigned. The sum collapes to $\frac{N!}{2}$ terms and acquires a 2 out front; that is, the prefactor becomes $\frac{2}{\sqrt{N!}}$. But this is wrong; if there are $\frac{N!}{2}$ distinct terms, the normalization should $\sqrt{\frac{2}{N!}}$. One simply has to correct the normalization for the fact that terms in the sum repeat and so the original normalization calculation that gave $\frac{1}{\sqrt{N!}}$ was wrong.

Indistinguishable Particles (cont.)

For fermions, the antisymmetrized state can be written conveniently as a determinant, usually termed a Slater Determinant:

$$|\psi_{n_1}, \cdots, \psi_{n_N}\rangle_- = \frac{1}{\sqrt{N!}} \begin{vmatrix} |\psi_{n_1}\rangle^{(1)} & \cdots & |\psi_{n_N}\rangle^{(1)} \\ \vdots & \ddots & \vdots \\ |\psi_{n_1}\rangle^{(N)} & \cdots & |\psi_{n_N}\rangle^{(N)} \end{vmatrix}$$

where the product between states when the determinant is calculated is taken to be the direct product, $\otimes$. Note how the particle number is fixed for each row (the superscript number) while the state is fixed for each column (the subscript): this structure gives the approprate result.

Indistinguishable Particles (cont.)

Let's consider a possible point of confusion on whether different ways of obtaining the same permutation are consistent. That is, it is quite important for fermions that the "oddness" or "evenness" of a given permutation be well-defined — that there not be a way of obtaining a given permutation by two different sets of pair exchanges, one with an odd number of exchanges and one with an even number — otherwise it will not be clear whether to assign a given permutation a negative sign or not.

This proof must also be done by induction. Suppose it is true for N particles that a given permutation requires either an odd or an even number of pairwise exchanges but never both. Now, tack on another particle and consider all the possible new permutations. There are two kinds of new permutations.

In the first kind, the $N+1$ particle remains in the $N+1$ spot. In this case, we just have a permutation of N particles, with the new particle tacked on to the end, so our assumption that permutations are either odd or even but never both applies — this new permutation of $N+1$ particles is also either odd or even but not both.

Indistinguishable Particles (cont.)

The second kind of permutation is the kind for which the $N+1$ particle is in position $j \leq N$ and the kth particle in the $N+1$ position ($j \neq k$ in general though $j = k$ is allowed). Call this permutation Π_1. Clearly, one can do a single pairwise exchange of the j and $N+1$ positions to put the $N+1$ particle in the $N+1$ position and the k particle in the j position. Now one has a permutation of N particles with the $N+1$ particle stuck on the end; call this permutation Π_2. So, whatever the sign of the permutation of N particles that yields Π_2, the original permutation of $N+1$ particles Π_1 is opposite in sign because it requires one more pairwise exchange. To complete the proof, we need to show that every way of obtaining Π_1 from Π_2 requires an odd number of pair exchanges; that is, we've shown that the sign of Π_2 is unique, but we need to show that the sign flip in getting from Π_2 to Π_1 is unique, also. Suppose it were not; suppose there were a path from Π_2 to Π_1 requiring an even number of exchanges; call this path B, and path A is our single pairwise exchange of particle k and particle $N+1$. Consider the path that consists of path B with the path A single pairwise exchange added on. This path has an odd number of exchanges, yet it takes the permutation Π_2 and returns the permutation Π_2. Π_2 is really a permutation of N particles only because the $N+1$ particle is stuck on the end, so this says that there are two paths, one odd and one even, of obtaining a permutation of N particles, contradicting the assumption that the uniqueness of the sign of permutations is true for N particles.

Proving the uniqueness of permutation sign for $N = 2$ particles is straightforward: there are only two permutations, one even and one odd, and any other path to those permutations requires doing an even number of unnecessary exchanges of the two particles, so the oddness or evenness of a permutation is unique.

Indistinguishable Particles (cont.)

Let's consider an implication of the above for non-adjacent and adjacent pairwise exchanges. An adjacent pairwise exchange is one in which you exchange particles that are adjacent to each other. A non-adjacent pairwise exchange is one in which two particles that are not adjacent to each other are exchanged. Our proof implies that the sign of a permutation arrived at by a single non-adjacent pairwise is consistent with the sign one would obtain by obtaining that permutation by a set of adjacent pairwise exchanges. To illustrate this, consider a 3-particle example. Consider the two permutations

$$1\ 2\ 3 \qquad 3\ 2\ 1$$

One can obtain the second from the first by the non-adjacent pairwise exchange $1 \leftrightarrow 3$, indicating it is an odd permutation. Is the path between these permutations by adjacent pairwise exchanges also odd? Yes:

$$1\ 2\ 3 \ \rightarrow\ 1\ 3\ 2 \ \rightarrow\ 3\ 1\ 2 \ \rightarrow\ 3\ 2\ 1$$

which requires three pairwise exchanges, also an odd number.

Indistinguishable Particles (cont.)

When to Symmetrize or Antisymmetrize

In some limit, we must be able to ignore the need to symmetrize or antisymmetrize states of a system of indistinguishable particles — the behavior of the electron in a hydrogen atom in Santa Monica does not care about the electrons in hydrogen atoms in a distant location such as, or example, Orange County. How do we recover this limit from the above formalism?

It is easy. Let's consider two indistinguishable particles that are created in a state

$$|\psi_a, \psi_b\rangle_{\pm}^{(1)\otimes(2)} = \frac{1}{\sqrt{2}}\left[|\psi_a\rangle^{(1)} \otimes |\psi_b\rangle^{(2)} \pm |\psi_b\rangle^{(1)} \otimes |\psi_a\rangle^{(2)}\right]$$

where $|\psi_a\rangle$ is localized in Santa Monica (preferably, in a cafe with a good view of the ocean and a good beer), and $|\psi_b\rangle$ is localized in Orange County (it doesn't really matter where because, to first order, every point in Orange County looks like every other point in Orange County). The probability of finding one particle at position x_{SM} in the cafe in Santa Monica and the other particle at position x_{OC} somewhere in Orange County is, by analogy to Equation 11.8,

$$\begin{aligned} P_{\pm}(x_{SM}, x_{OC}) = {} & |\psi_{a,x}(x_{SM})|^2 |\psi_{b,x}(x_{OC})|^2 + |\psi_{b,x}(x_{SM})|^2 |\psi_{a,x}(x_{OC})|^2 \\ & \pm \psi_{a,x}^*(x_{SM})\, \psi_{b,x}(x_{SM})\, \psi_{b,x}^*(x_{OC})\, \psi_{a,x}(x_{OC}) \\ & \pm \psi_{b,x}^*(x_{SM})\, \psi_{a,x}(x_{SM})\, \psi_{a,x}^*(x_{OC})\, \psi_{b,x}(x_{OC}) \end{aligned}$$

Indistinguishable Particles (cont.)

Now, let's calculate the probability of finding one of the particles at a position x_{SM} somewhere in that cafe without caring where the second particle is. We obtain that probability by integrating over x_{OC}: any outcome for x_{OC} is fine for our purposes. That is:

$$\begin{aligned} P_{\pm}(x_{SM}) = |\psi_{a,x}(x_{SM})|^2 \int_{-\infty}^{\infty} dx_{OC}\, |\psi_{b,x}(x_{OC})|^2 \\ + |\psi_{b,x}(x_{SM})|^2 \int_{-\infty}^{\infty} dx_{OC}\, |\psi_{a,x}(x_{OC})|^2 \\ \pm \psi^*_{a,x}(x_{SM})\, \psi_{b,x}(x_{SM}) \int_{-\infty}^{\infty} dx_{OC}\, \psi^*_{b,x}(x_{OC})\, \psi_{a,x}(x_{OC}) \\ \pm \psi^*_{b,x}(x_{SM})\, \psi_{a,x}(x_{SM}) \int_{-\infty}^{\infty} dx_{OC}\, \psi^*_{a,x}(x_{OC})\, \psi_{b,x}(x_{OC}) \end{aligned}$$

We may neglect the last three terms: they all contain at least one power of $\psi_{b,x}(x_{SM})$, which is vanishing according to our localization assumptions for $|\psi_a\rangle$ and $|\psi_b\rangle$. The integral in the first term gives unity because the single-particle states are assumed to be individually normalized.

Indistinguishable Particles (cont.)

So we have

$$P_{\pm}(x_{SM}) = |\psi_{a,x}(x_{SM})|^2$$

which is what we would have obtained had we neglected the other particle. As one might expect, indistinguishability only matters when particles are close enough that their wavefunctions have some overlap.

Section 12

Symmetries

Lecture 35:
Coordinate Transformations:
Passive Coordinate Transformations

Date Revised: 2009/01/23
Date Given: 2009/01/21

Coordinate Transformations

A coordinate transformation is any relabeling of the points in space from set of axes F with coordinates q to a different set of axes F' with coordinates q', with $q' = q'(q, t)$. Some examples:

- Mirror transformation:

$$x' = -x \qquad y' = y \qquad z' = z$$

- Rotation transformation about the z axis by θ (CCW):

$$x' = x\cos\theta + y\sin\theta \qquad y' = -x\sin\theta + y\cos\theta \qquad z' = z$$

- Translation:

$$x' = x - a \qquad y' = y \qquad z' = z$$

In all of these, a new set of axes are being imposed on space and one chooses to label points relative to the new system rather than the old one. You should draw the old and new axes so you are clear on what is meant by each transformation.

Discrete vs. Continuous Transformations

A discrete coordinate transformation is one for which there is no continuous parameter that varies the effect of the transformation. A continuous coordinate transformation is therefore one that does depend on a continuous parameter. Continuous transformations can be differentiated with respect to the parameter.

Examples of discrete transformations include parity transformation (inversion of the coordinates through the origin) and mirror transformations (inversion through a plane; of course, one can vary the plane, but once one picks a plane, there is just one transformation) are a couple obvious ones.

Continuous transformations include spatial translations and rotations, where, even after one has picked a direction for the transformation (a translation direction or rotation axis), there is a continuous parameter: the translation distance or rotation angle in these cases.

Discrete transformations in general need to be treated on a case-by-case basis, whereas continuous transformation can be treated in a unified manner. The latter also result in conserved quantities, as we shall see.

Passive Coordinate Transformations (cont.)

Effect of Coordinate Transformations in Quantum Mechanics

A relabeling of the above type has no effect on the particle state $|\psi\rangle$ because it is an abstract object and makes no reference to a particular coordinate system. But, there are two ways in which the relabeling matters:

- First, because of the redefinition of the coordinate axes, the position and momentum basis change; there are new sets of basis states $\{|q'\rangle\}$ and $\{|p'_q\rangle\}$ that indicate states with well-defined position or momentum relative to the new axes rather than the old axes. The old position and momentum bases $\{|q\rangle\}$ and $\{|p_q\rangle\}$ still exist and the projections $\langle q|\psi\rangle$ and $\langle p_q|\psi\rangle$ are unchanged.
- Second, it affects the operators. We want the above new basis elements to be eigenvectors of position and momentum operators, so we require new position and momentum operators $\{Q'\}$ and $\{P'_q\}$, as these new basis elements will not be eigenvectors of the original position and momentum operators with the expected eigenvalues. These are not the same as the old operators $\{Q\}$ and $\{P_q\}$, meaning that the action of $\{Q'\}$ and $\{P'_q\}$ on states are different from that of $\{Q\}$ and $\{P_q\}$. The $\{Q\}$ and $\{P_q\}$ operators of course still exist and there is no change in their action.

Realize that the new bases live in the same Hilbert space as the old bases, and that the new and old operators both act there; the coordinate relabeling results in no new Hilbert space!

Passive Coordinate Transformations

We define the passive transformation operator T_P to be the operator that maps from the original basis $\{|q\rangle\}$ to the new basis $\{|q'\rangle\}$, element-by-element:

$$e^{i\theta}|q'\rangle = T_P\,|q\rangle$$

where θ is real. That is, the passive transformation operator "transforms" the position-basis elements associated with the old coordinate system into position-basis elements associated with the new coordinate system. (We will define an *active* transformation below.) The $e^{i\theta}$ factor is allowed out front because it does not affect the norm of the state or the unitarity property derived below. In general, we will choose $\theta = 0$ so that the factor's value is always unity, but that is an additional assumption that must be made; it need not always hold true.

Since we are just relabeling points in space, the dimensionality of the new basis must be the same as that of the old basis; that is, the mapping is one-to-one, or invertible. Assuming the new basis elements are normalized in the same way as the old ones (which is unaffected by the unity modulus factor), the transformation also preserves inner products between basis elements because the unit-modulus factor cancels out. Therefore, it preserves all inner products, and hence is unitary, $T_P^\dagger = T_P^{-1}$.

Passive Coordinate Transformations (cont.)

We **define** position and momentum operators in the new basis by the requirement that their action in the new basis be the same as the corresponding old operators' action in the old basis; this follows from their being position and momentum operators for the new coordinate system. So,

$$\langle q_1' \,|Q'|q_2' \rangle = \langle q_1 \,|Q\,|q_2 \rangle$$

By the definition of the transformation operator's action on states, $|q' \rangle = T_P|q \rangle$, we also have $\langle q_1' \,|Q'|q_2' \rangle = \langle q_1 \,|T_P^\dagger Q' T_P|q_2 \rangle$. Combining the two statements gives

$$\langle q_1 \,|Q\,|q_2 \rangle = \langle q_1 \,|T_P^\dagger Q' T_P|q_2 \rangle$$

Since this relation holds for all $|q_1 \rangle$ and $|q_2 \rangle$, it therefore holds that

$$Q' = T_P\, Q\, T_P^\dagger = T_P\, Q\, T_P^{-1}$$

The above proof carries through for the $\{P_q\}$ also, and, in fact, for any operator. Because it holds for any pair of basis elements, it holds for any pair of states, giving

$$\langle v' \,|O'|w' \rangle = \langle v \,|O|w \rangle \qquad \Longleftrightarrow \qquad O' = T_P Q T_P^\dagger \tag{12.1}$$

Passive Coordinate Transformations (cont.)

We may thus think of the passive transformation operator as not only transforming the basis elements but also transforming any operator, with the new operator **defined** by the requirement that the action of the new operator in the new basis is the same as the old operator in the old basis.

Note that the passive transformation operator affects any basis, not just the position basis. All of our bases are defined by being eigenbases of some Hermitian operator (Q, P_q, etc.). If the transformation changes these operators, then their eigenbases will change, too. One obtains the transformed eigenbases by diagonalizing the transformed operators.

We note that, in general, in addition to the defining transformation relation Equation 12.1, it is also possible to rewrite the primed operators in terms of the unprimed operators simply because the primed observables are also observables in the unprimed coordinate system and hence must be functions of the $\{Q\}$ and $\{P_q\}$. The form of these latter relations will depend on the particular type of transformation, though.

Additionally, because we do not want to move the potentials and particles when we perform a coordinate transformation, we do not in general want to transform H to H'. We will, though, write H in terms of the $\{Q'\}$ and $\{P_q'\}$ operators by using this latter set of relations that write the primed operators in terms of the unprimed operators (and vice versa).

Example 12.1: Passive Mirror Transformation of a Particle in a Two-Dimensional Box

Consider a particle in the an eigenstate of a two-dimensional box of dimension $L_1 \times L_2$ whose corner is at the origin. Let the state be denoted by $|\psi_{ab}\rangle$. The energy is

$$E_{ab} = \frac{\hbar^2 \pi^2}{2m}\left(\frac{a^2}{L_1^2} + \frac{b^2}{L_2^2}\right)$$

The position-basis representation is

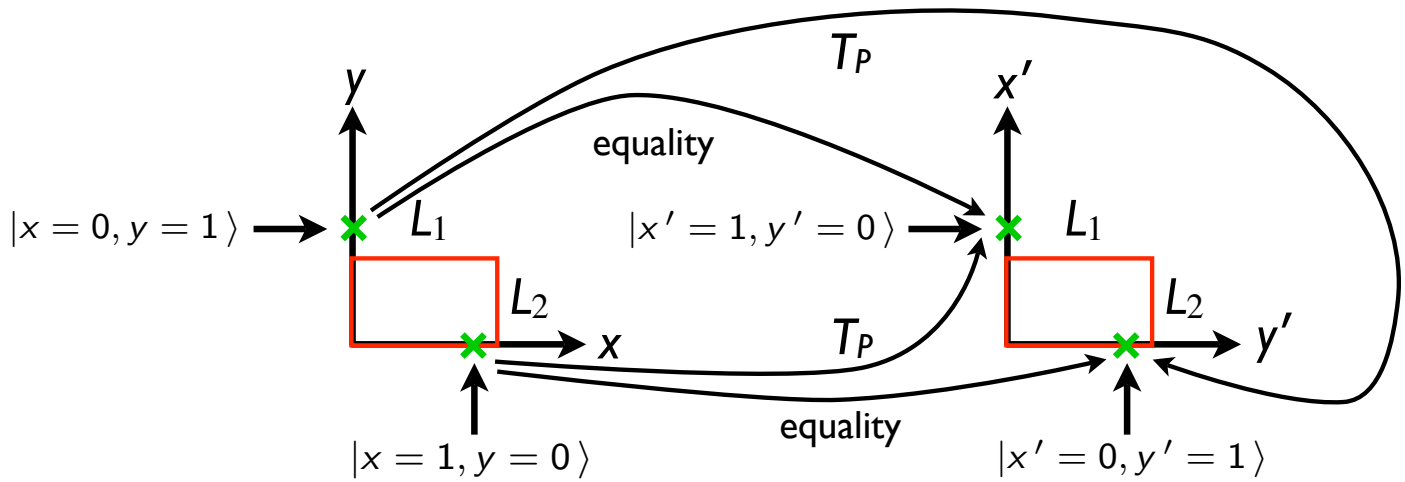
$$\psi_{ab}(x,y) = \langle x, y | \psi_{ab}\rangle = \sqrt{\frac{4}{L_1 L_2}} \sin\left(\frac{a\pi x}{L_1}\right) \sin\left(\frac{b\pi y}{L_2}\right)$$

(Note that, because the box's corner, not center, is at the origin, the sines are the only allowed states. We do this to avoid having to carry around sines and cosines in this example.) and we would find the expectation values

$$\langle \psi_{ab}|X|\psi_{ab}\rangle = \frac{L_1}{2} \quad \langle \psi_{ab}|Y|\psi_{ab}\rangle = \frac{L_2}{2} \quad \langle \psi_{ab}|P_x|\psi_{ab}\rangle = 0 \quad \langle \psi_{ab}|P_y|\psi_{ab}\rangle = 0$$

Passive Coordinate Transformations (cont.)

Now, perform the coordinate transformation $x' = y, y' = x$, which is a mirroring through the line $x = y$. The transformation is visualized in the figure below. We see clearly that the axes are relabeled but the potential has not moved in space; this is what we mean by a passive transformation.



Passive Coordinate Transformations (cont.)

The transformation gives $|x' = u, y' = v\rangle = T_P|x = u, y = v\rangle$ where u and v are just numbers, not tied to any coordinate system. We'll see below why we use the u and v. Be sure you understand what is meant by this notation, in particular the ordering of u and v: for example, a state along the x axis, $|x = 1, y = 0\rangle$, gets mapped to a state along the x' axis, $|x' = 1, y' = 0\rangle$. The transformation rules tell us that $|x' = u, y' = v\rangle = |x = v, y = u\rangle$. That is, when written in terms of unprimed coordinates, the state that $|x = u, y = v\rangle$ gets mapped to by T_P is the state $|x = v, y = u\rangle$ in the unprimed coordinates. In the example, the state on the x' axis $|x' = 1, y' = 0\rangle$ is equal to the state on the y axis $|x = 0, y = 1\rangle$, and the state $|x' = 0, y' = 1\rangle$ on the y' axis is equal to the state $|x = 1, y = 0\rangle$ on the x axis. Make sure you have the above straight; it *is* confusing! These relationships are also indicated in the figure.

The state ψ_{ab} stays unchanged, but it must now be written in terms of the basis elements in the new axes, the $\{|x', y'\rangle\}$. We denote this wavefunction by $\psi_{ab,q'}(x', y')$ and it is given by $\psi_{ab,q'}(x', y') = \langle x', y'|\psi_{ab}\rangle$. We need the extra $_{q'}$ subscript to indicate that $\psi_{ab,q'}$ is a different function of its arguments than the old $\psi_{ab}(x, y)$, which we shall now denote by $\psi_{ab,q}(x, y)$. This is consistent with our generic notation of denoting a wavefunction in some particular representation (basis) with a subscript that specifies the representation: the $\{|x', y'\rangle\}$ and $\{|x, y\rangle\}$ are different bases, so the representation of $|\psi_{ab}\rangle$ in the two bases are different and need to be distinguished.

Passive Coordinate Transformations (cont.)

How do we calculate $\psi_{ab,q'}(x',y')$, explicitly? We recall that $|x'=u,y'=v\rangle = |x=v,y=u\rangle$. So, we have:

$$\begin{aligned}\psi_{ab,q'}(x'=u,y'=v) &= \langle x'=u,y'=v\,|\psi_{ab}\rangle \\ &= \langle x=v,y=u\,|\psi_{ab}\rangle \\ &= \psi_{ab,q}(x=v,y=u) \\ &= \sqrt{\frac{4}{L_1L_2}}\,\sin\left(\frac{a\,\pi\,v}{L_1}\right)\,\sin\left(\frac{b\,\pi\,u}{L_2}\right)\end{aligned}$$

$$\text{So,}\qquad \psi_{ab,q'}(x',y') = \sqrt{\frac{4}{L_1L_2}}\,\sin\left(\frac{b\,\pi\,x'}{L_2}\right)\,\sin\left(\frac{a\,\pi\,y'}{L_1}\right)$$

We went through these specific numbers u and v to avoid confusion about when you replace (x,y) by (x',y') and when by (y',x'): u and v are just numbers, not tied to any coordinate system. We know that the state at $|x'=u,y'=v\rangle$ is the same state as $|x=v,y=u\rangle$, and knowing that lets us rewrite the wavefunction in terms of x' and y'. The above form should be intuitively obvious, though; if the x' axis lies along the y axis and the y' axis lies along the x axis, and the state does not change, then x' must take on the role of y and y' of x.

Passive Coordinate Transformations (cont.)

What are the new operators? As we said, the formula $O' = T_P\, O\, T_P^\dagger$ is abstract. To figure it out, we resort to matrix elements. We require

$$\langle x_1', y_1' \,|X'|x_2', y_2' \rangle = \langle x_1, y_1 \,|X\,|x_2, y_2 \rangle$$

Again, it will be far clearer if we use specific numbers not tied to a coordinate system. So let's write the above as

$$\langle x' = u_1, y' = v_1 \,|X'|x' = u_2, y' = v_2 \rangle = \langle x = u_1, y = v_1 \,|X\,|x_2 = u_2, y_2 = v_2 \rangle$$

Note how the u's and v's are distributed: the state $|x' = u_1, y' = v_1 \rangle$ satisfies $|x' = u_1, y' = v_1 \rangle = T_P|x = u_1, y = v_1 \rangle$ and also $|x_1' = u_1, y_1' = v_1 \rangle = |x = v_1, y = u_1 \rangle$. The matrix element on the right is

$$\langle x = u_1, y = v_1 \,|X\,|x = u_2, y = v_2 \rangle = u_1\, \delta(u_1 - u_2)\, \delta(v_1 - v_2)$$

Passive Coordinate Transformations (cont.)

So our requirement becomes

$$\langle x' = u_1, y' = v_1 \,|X'|x' = u_2, y' = v_2 \,\rangle = u_1\,\delta(u_1 - u_2)\,\delta(v_1 - v_2)$$

which may now be rewritten as

$$\langle x', y' \,|X'|x', y' \,\rangle = x_1'\,\delta(x_1' - x_2')\,\delta(y_1' - y_2')$$

Thus, our operator X' does exactly what we expect it to do: it has as eigenvectors the states $\{|x', y' \,\rangle\}$ with eigenvalues $\{x'\}$. It therefore is the X' operator.

If you do not use the u's and v's when trying to prove the above statement, it is very easy to get confused and end up thinking that the X' operator pulls out the y' coordinate of $|x', y' \,\rangle$ because you end up with x_1 instead of u_1 at the front of the expression, and the transformation says $y_1' = x_1$, so one is tempted to replace the x_1 with y_1' and think that one gets Y' instead of X'. This problem occurs because it is unclear whether one should use the relation $(x_1, y_1) \rightarrow (x_1', y_1')$ or the relation $|x_1', y_1' \,\rangle = |y_1, x_1 \,\rangle$. By instead using these "numbers" (u_1, v_1) and (u_2, v_2), one doesn't get into notational quandaries like this.

Passive Coordinate Transformations (cont.)

By analogy, we may also conclude

$$\langle x_1', y_1' | Y' | x_2', y_2' \rangle = y_1' \, \delta(x_1' - x_2') \, \delta(y_1' - y_2')$$

$$\langle x_1', y_1' | P_x' | x_2', y_2' \rangle = -i\hbar \frac{d}{dx_1'} \delta(x_1' - x_2') \, \delta(y_1' - y_2')$$

$$\langle x_1', y_1' | P_y' | x_2', y_2' \rangle = -i\hbar \frac{d}{dy_1'} \delta(x_1' - x_2') \, \delta(y_1' - y_2')$$

That is, all the primed operators do exactly what we expect them to in terms of the primed coordinate system. In obtaining the above, we will have taken a derivative with respect to the "numbers" u_1 and v_1 because they are indeed still variables: the point of introducing them is that they are not notationally tied to a particular coordinate system, so they remove the confusion noted above.

We have defined the primed operators by requiring that their matrix elements in the primed basis be the same as the matrix elements of the unprimed operators between the unprimed basis elements that are transformed into the primed basis elements. Since the unprimed and primed basis elements are not equal, this is **not** a statement of equality between the unprimed operators and primed operators.

Passive Coordinate Transformations (cont.)

However, because the unprimed and primed basis elements describe the same Hilbert space and there are equality relations between them (as illustrated in the figure), we can calculate the matrix elements of the primed operators in the unprimed basis and compare these to the matrix elements of the unprimed operators in the unprimed basis to obtain a relation between the primed and unprimed operators (or conversely, we could calculate the matrix elements of the unprimed operators in the primed basis and compare to the matrix elements of the primed operators in the primed basis).

Explicitly:

$$\begin{aligned}
\langle x = u_1, y = v_1 \,|X'|x = u_2, y = v_2 \,\rangle &= \langle x' = v_1, y' = u_1 \,|X'|x' = v_2, y' = u_2 \,\rangle \\
&= v_2 \langle x' = v_1, y' = u_1 \,|x' = v_2, y' = u_2 \,\rangle \\
&= v_2 \langle x = u_1, y = v_1 \,|x = u_2, y = v_2 \,\rangle \\
&= \langle x = u_1, y = v_1 \,|Y \,|x = u_2, y = v_2 \,\rangle \\
\Longrightarrow \quad \langle x_1, y_1 \,|X'|x_2, y_2 \,\rangle &= \langle x_1, y_1 \,|Y \,|x_2, y_2 \,\rangle
\end{aligned}$$

We find that all the possible matrix elements between unprimed basis elements of X' and Y are equal, so they must be the same operator:

$$X' = Y$$

.

Passive Coordinate Transformations (cont.)

But we also have a relation between X' and X. Let's write both of these down:

$$X' = T_P X\, T_P^\dagger \qquad\qquad X' = Y$$

Interesting, eh? The former is true, regardless of the situation, but the latter is far more useful in trying to understand this specific example. Similar relations hold for Y', P_x', and P_y':

$$\begin{aligned} Y' &= T_P Y\, T_P^\dagger & Y' &= X \\ P_x' &= T_P P_x T_P^\dagger & P_x' &= P_y \\ P_y' &= T_P P_y T_P^\dagger & P_y' &= P_x \end{aligned}$$

For P_x' and P_y', one establishes the relations in the right column by taking matrix elements in the momentum basis rather than the position basis because these operators are diagonal in the momentum basis but not in the position basis. One can of course establish relationships between primed and unprimed momentum basis elements as we have done for position basis elements.

Simple relations like these do not always hold: they are specific to our mirror transformation example. As we noted above, though, it is always possible to write the $\{Q'\}$ and $\{P_{q'}'\}$ operators in terms of the $\{Q\}$ and $\{P_q\}$ operators.

Passive Coordinate Transformations (cont.)

As we explained above, we do not transform the Hamiltonian, we simply rewrite it in terms of primed operators. But, for the sake of pedagogy, let us explicitly transform the Hamiltonian and see what goes wrong:

$$H' = T_P \left(\frac{P_x^2 + P_y^2}{2\,m} + V(X, Y) \right) T_P^\dagger = \frac{(P_x')^2 + (P_y')^2}{2\,m} + V(X', Y')$$

where we can make the step by thinking of V in terms of a Taylor expansion and inserting $T_P^\dagger\, T$ between each factor of X, Y, P_x, or P_y. Nothing obviously wrong happens in the kinetic terms, but the potential term has X' as its first argument and Y' as its second argument: the L_1 dimension of the box is along x' and the L_2 dimension along y'. This corresponds to mirroring the box, which we did not want to do!

With that out of the way, let us rewrite H in terms of X', Y', P_x', and P_y' using the relations like $X' = Y$ proved above:

$$H = \frac{P_x^2 + P_y^2}{2\,m} + V(X, Y) = \frac{(P_y')^2 + (P_x')^2}{2\,m} + V(Y', X')$$

X' and Y' are in the right arguments of V to have the L_1 side of the box along y' and the L_2 side along x'.

Passive Coordinate Transformations (cont.)

The above rewriting of H tells us two things: first, our prior rewriting of the eigenstate wavefunctions in the $\{|x', y'\rangle\}$ basis was correct: we expect the x' coordinate to be linked up with L_2 and the y' coordinate to be linked up with L_1, as we found earlier for $\psi_{ab,q'}(x', y')$. Second, the energies of the states are unchanged because the Hamiltonian is unchanged.

We may calculate the expectation values of the operators in two ways. The first method is to do the necessary integrals for, *e.g.*, $\langle\psi_{ab}|X'|\psi_{ab}\rangle$, explicitly by inserting completeness relations for the new $\{|x', y'\rangle\}$ basis. The second method is to simply use the relations between the primed and unprimed operators (like $X' = Y$) that we proved above, making the whole exercise trivial:

$$\langle\psi_{ab}|X'|\psi_{ab}\rangle = \langle\psi_{ab}|Y|\psi_{ab}\rangle = \frac{L_2}{2}$$
$$\langle\psi_{ab}|Y'|\psi_{ab}\rangle = \langle\psi_{ab}|X|\psi_{ab}\rangle = \frac{L_1}{2}$$
$$\langle\psi_{ab}|P_x'|\psi_{ab}\rangle = \langle\psi_{ab}|P_y|\psi_{ab}\rangle = 0$$
$$\langle\psi_{ab}|P_y'|\psi_{ab}\rangle = \langle\psi_{ab}|P_x|\psi_{ab}\rangle = 0$$

Lecture 36:
Coordinate Transformations:
Passive Coordinate Transformations

Date Revised: 2009/03/13
Date Given: 2009/01/23

Example 12.2: Passive Rotation Transformation of a Particle in a Two-Dimensional Box

Let's consider the same Hamiltonian as in the previous example, but let the coordinate transformation be rotation in the plane by an angle θ (CCW when viewed from above):

$$x' = x\,\cos\theta + y\,\sin\theta \qquad\qquad y' = -x\,\sin\theta + y\,\cos\theta$$

For the sake of brevity, we will write the above as

$$x' = x\,c_\theta + y\,s_\theta \qquad\qquad y' = -x\,s_\theta + y\,c_\theta$$

As before, let's first understand the relation between elements of the unprimed and primed bases. Our generic transformation rule is

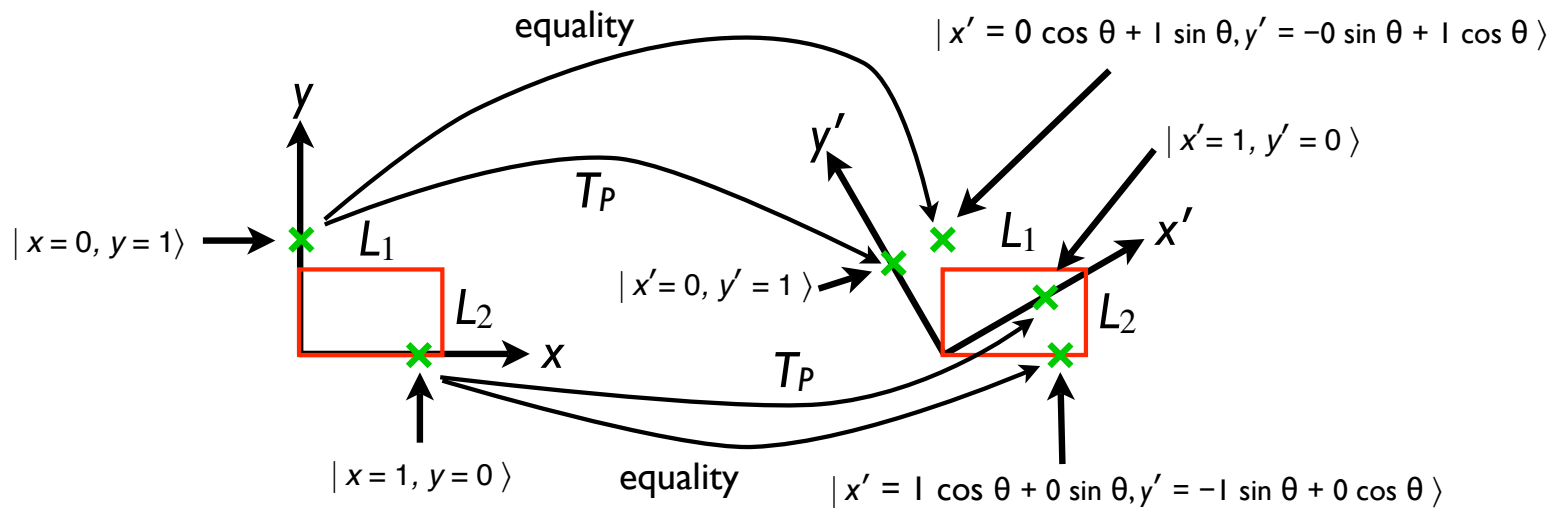
$$|x' = u, y' = v\rangle = T_P|x = u, y = v\rangle$$

Passive Coordinate Transformations (cont.)

Because the primed and unprimed basis elements describes states in the same Hilbert space, and using the following figure, we recognize that the following equalities hold also:

$$|x = u\,c_\theta + v\,s_\theta, y' = -u\,s_\theta + v\,c_\theta\rangle = |x = u, y = v\rangle$$
$$|x' = u, y' = v\rangle = |x = u\,c_\theta - v\,s_\theta, y = u\,s_\theta + v\,c_\theta\rangle$$

The above relationships are illustrated in the following figure, where we consider the specific basis elements $|x = 1, y = 0\rangle$ and $|x = 0, y = 1\rangle$.



Passive Coordinate Transformations (cont.)

Following along our previous example, let's now calculate the primed position basis wavefunction using the expressions obtained above that related unprimed and primed basis elements:

$$\begin{aligned}\psi_{ab,q'}(x'=u,y'=v) &= \langle x'=u,y'=v\,|\psi_{ab}\,\rangle \\ &= \langle x=u\,c_\theta - v\,s_\theta, y=u\,s_\theta + v\,c_\theta\,|\psi\,\rangle \\ &= \sqrt{\frac{4}{L_1 L_2}}\sin\left(\frac{a\,\pi\,(u\,c_\theta - v\,s_\theta)}{L_1}\right)\sin\left(\frac{b\,\pi\,(u\,s_\theta + v\,c_\theta)}{L_2}\right)\end{aligned}$$

$$\text{So, } \psi_{ab,q'}(x',y') = \sqrt{\frac{4}{L_1 L_2}}\sin\left(\frac{a\,\pi\,(x'c_\theta - y's_\theta)}{L_1}\right)\sin\left(\frac{b\,\pi\,(x's_\theta + y'c_\theta)}{L_2}\right)$$

The form of the wavefunction in the primed coordinates is of course more complicated because the box is not aligned with the primed coordinate axes.

Passive Coordinate Transformations (cont.)

Next, let's use the relationship between basis elements to obtain the matrix elements of the new operators X', Y', P_x', and P_y', again along the lines of what we did in the previous example. Because of the defining relationship in Equation 12.1, $\langle q_1' | O' | q_2' \rangle = \langle q_1 | O | q_2 \rangle$, we may write

$$\begin{aligned}\langle x' = u_1, y' = v_1 | X' | x' = u_2, y' = v_2 \rangle &= \langle x = u_1, y = v_1 | X | x = u_2, y = v_2 \rangle \\ &= u_1 \, \delta(u_1 - u_2) \, \delta(v_1 - v_2) \\ \Longrightarrow \quad \langle x_1', y_1' | X' | x_2', y_2' \rangle &= x_1' \delta(x_1' - x_2') \delta(y_1' - y_2')\end{aligned}$$

Analogously, we have

$$\begin{aligned}\langle x_1', y_1' | Y' | x_2', y_2' \rangle &= y_1' \delta(x_1' - x_2') \delta(y_1' - y_2') \\ \langle x_1', y_1' | P_x' | x_2', y_2' \rangle &= -i\,\hbar \frac{d}{dx_1'} \, \delta(x_1' - x_2') \, \delta(y_1' - y_2') \\ \langle x_1', y_1' | P_y' | x_2', y_2' \rangle &= -i\,\hbar \frac{d}{dy_1'} \, \delta(x_1' - x_2') \, \delta(y_1' - y_2')\end{aligned}$$

You will notice that these kinds of relations are always the same for *any* coordinate transformation. That is the whole idea of Equation 12.1: we want the action of the primed operators in the primed basis to be the same as the action of the unprimed operators in the unprimed basis.

Passive Coordinate Transformations (cont.)

Let's next use the matrix elements to obtain an explicit relationship between the primed and unprimed operators. For X', we have

$$
\begin{aligned}
&\langle x = u_1, y = v_1 \,|X'|x = u_2, y = v_2 \,\rangle \\
&= \langle x' = u_1 c_\theta + v_1 s_\theta, y = -u_1 s_\theta + v_1 c_\theta \,|X'|x' = u_2 c_\theta + v_2 s_\theta, y' = -u_2 s_\theta + v_2 c_\theta \,\rangle \\
&= (u_2 c_\theta + v_2 s_\theta) \\
&\qquad \times \langle x' = u_1 c_\theta + v_1 s_\theta, y' = -u_1 s_\theta + v_1 c_\theta \,|x' = u_2 c_\theta + v_2 s_\theta, y' = -u_2 s_\theta + v_2 c_\theta \,\rangle \\
&= (u_2 c_\theta + v_2 s_\theta) \,\langle x = u_1, y = v_1 \,|x = u_2, y = v_2 \,\rangle \\
&= \langle x = u_1, y = v_1 \,|\, (X\, c_\theta + Y\, s_\theta) \,|x = u_2, y = v_2 \,\rangle \\
\Longrightarrow \qquad &\langle x_1, y_1 \,|X'|x_2, y_2 \,\rangle = \langle x_1, y_1 \,|\, (X\, c_\theta + Y\, s_\theta) \,|x_2, y_2 \,\rangle \\
\Longrightarrow \qquad &X' = X\, c_\theta + Y\, s_\theta
\end{aligned}
$$

Similarly, $\quad Y' = -X\, s_\theta + Y\, c_\theta \qquad P_x' = P_x c_\theta + P_y\, s_\theta \qquad P_y' = -P_x\, s_\theta + P_y c_\theta$

For P_x' and P_y', one establishes the above relations by taking matrix elements in the momentum basis rather than the position bsis because these operators are diagonal in the momentum basis but not in the position basis. One can of course establish relationships between primed and unprimed momentum basis elements as we have done for position basis elements.

Passive Coordinate Transformations (cont.)

Finally, let's rewrite the original Hamiltonian in terms of the primed operators as we did for the previous example:

$$\begin{aligned} H &= \frac{P_x^2 + P_y^2}{2\,m} + V(X, Y) \\ &= \frac{\left(P_x' c_\theta - P_y' s_\theta\right)^2 + \left(P_x' s_\theta + P_y' c_\theta\right)^2}{2\,m} + V\left(X' c_\theta - Y' s_\theta, X' s_\theta + Y' c_\theta\right) \\ &= \frac{\left[P_x'\right]^2 + \left[P_y'\right]^2}{2\,m} + V\left(X' c_\theta - Y' s_\theta, X' s_\theta + Y' c_\theta\right) \end{aligned}$$

The kinetic term remains simple in the new coordinate system, but the potential term becomes complicated, which is expected because the box is not aligned with the primed axes.

Lecture 37:
Coordinate Transformations:
Continuous Passive Coordinate Transformations
Active Coordinate Transformations

Date Revised: 2009/01/28

Date Given: 2009/01/26

General Properties of Continuous Passive Coordinate Transformations

Consider a passive coordinate transformation that can be parameterized by and differentiated with respect to a continuous variable ε, $T_P = T_P(\varepsilon)$. This might be the translation vector for a translation, the rotation angle for a rotation (such as in the example we just did), etc. Are there any interesting properties of such transformations that arise from this additional differentiability property?

The first point we may make is that the infinitesimal version of such a transformation (that is, $T_P(\varepsilon)$ for $\varepsilon \to 0$) may always be written in the form

$$T_P(\varepsilon) = I - \frac{i}{\hbar}\,\varepsilon\, G$$

where G is some operator to be determined. This property arises simply because $T_P(\varepsilon) \to I$ as $\varepsilon \to 0$. The choice of the $-\frac{i}{\hbar}$ coefficient is of course motivated by prior knowledge of what will happen below, but is completely general.

Generators for Continuous Coordinate Transformations (cont.)

What may we say about G? Let's investigate the consequences of the unitarity of $T_P(\varepsilon)$. To do this, we need to know what $T_P^{-1}(\varepsilon)$ is. It is $T_P^{-1}(\varepsilon) = I + \frac{i}{\hbar}\,\varepsilon\,G$, which we can prove by applying it to $T_P(\varepsilon)$:

$$T_P^{-1}(\varepsilon)\,T_P(\varepsilon) = \left(I + \frac{i}{\hbar}\,\varepsilon\,G\right)\left(I - \frac{i}{\hbar}\,\varepsilon\,G\right) = I + \mathcal{O}(\varepsilon^2) = I$$

where we drop $\mathcal{O}(\varepsilon^2)$ terms because we are assuming ε is infinitesimal. From the above, we can see that G must be Hermitian:

$$\begin{aligned} T_P^{\dagger}(\varepsilon) &= T_P^{-1}(\varepsilon) \\ I^{\dagger} + \frac{i}{\hbar}\,G^{\dagger} &= I + \frac{i}{\hbar}\,\varepsilon\,G \\ G^{\dagger} &= G \end{aligned}$$

Generators for Continuous Coordinate Transformations (cont.)

With some knowledge of the properties of G, we may construct the full transformation for arbitary ε by taking an infinite product of infinitesimal transformations:

$$T_P(\varepsilon) = \lim_{N\to\infty}\left[T_P\left(\frac{\varepsilon}{N}\right)\right]^N = \lim_{N\to\infty}\left[I - \frac{i}{\hbar}\,\frac{\varepsilon}{N}\,G\right]^N = \exp\left(-\frac{i}{\hbar}\,\varepsilon\,G\right)$$

where ε/N is infinitesimal as $N\to\infty$. We have used a generic property of the exponential function in the last step, one that can be verified to be equivalent to the standard Taylor expansion via inductive proof. That G is Hermitian is now not surprising, as we know from previous work that $e^{i\Lambda}$ is unitary if Λ is Hermitian. The above formula leads us to an explicit formula for G:

$$T_P^{-1}(\varepsilon)\left[i\,\hbar\,\frac{d}{d\varepsilon}\,T_P(\varepsilon)\right] = T_P^{-1}(\varepsilon)\,T_P(\varepsilon)\,G = G$$

That is, G can be obtained by differentiating $T_P(\varepsilon)$. Due to the relation between G and $T_P(\varepsilon)$, G is called the generator of the passive coordinate transformation $T_P(\varepsilon)$. Because G is Hermitian, G is allowed to be an observable, a property that will be interesting for symmetry transformations.

Generators for Continuous Coordinate Transformations (cont.)

We note that the explicit form for $T_P(\varepsilon)$ for continuous transformations lets us write down a more explicit form for how position-basis elements, states, and operators transform:

$$|q'\rangle = \exp\left(-\frac{i}{\hbar}\varepsilon G\right)|q\rangle$$
$$O' = \exp\left(-\frac{i}{\hbar}\varepsilon G\right) O \exp\left(\frac{i}{\hbar}\varepsilon G\right)$$

That is, rather than specifying $T_P(\varepsilon)$ by the mapping from unprimed to primed basis elements, we can now specify it simply by the action of G.

The effect of the transformation $T_P(\varepsilon)$ with generator G may be particularly simply written in terms of the eigenstates of G. Consider first the action of the transformation on an eigenstate $|g\rangle$ of G:

$$|g'\rangle = e^{-\frac{i}{\hbar}\varepsilon G}|g\rangle = e^{-\frac{i}{\hbar}\varepsilon g}|g\rangle$$

That is, the eigenstates of G transform very simply under $T_P(\varepsilon)$: they retain the same direction in Hilbert space, picking up only a unity-modulus factor.

Generators for Continuous Coordinate Transformations (cont.)

The above suggests that the transformation of the $\{|g\rangle\}$-basis wavefunction of the state is trivial. This wavefunction is $\psi_g(g) = \langle g\,|\psi\rangle$. Then we have

$$\psi_{g'}(g') = \langle g'\,|\psi\rangle = \langle g\,|\,e^{\frac{i}{\hbar}\,\varepsilon\,G}\,|\psi\rangle = e^{\frac{i}{\hbar}\,\varepsilon\,g}\,\langle g\,|\psi\rangle = e^{\frac{i}{\hbar}\,\varepsilon\,g}\,\psi_g(g)$$

Note the sign in the argument of the expoential, which arises because we have $\langle g'\,|$ not $|g'\rangle$ in the expression. Our result shows that the $\{|g\rangle\}$-basis wavefunction simply picks up a g-dependent unity-modulus factor. Of course, the g dependence of that factor is what results in interesting behavior. But, clearly, things are much simplified in this form.

Finally, the above encourages us to look at the transformation of O in the $\{|g\rangle\}$ basis. That is:

$$\langle g_1\,|O'|g_2\rangle = \langle g_1\,|e^{-\frac{i}{\hbar}\,\varepsilon\,G}\,O\,e^{\frac{i}{\hbar}\,\varepsilon\,G}|g_2\rangle = e^{-\frac{i}{\hbar}\,\varepsilon\,(g_1-g_2)}\,\langle g_1\,|O\,|g_2\rangle$$

Again, we get a very simple relation between the matrix elements of O' and those of O in the $\{|g\rangle\}$ basis. (Note that the above is different from the transformed operator defining relation, Equation 12.1, which here takes the form $\langle g_1'\,|O'|g_2'\rangle = \langle g_1\,|O|g_2\rangle$.)

Explicitly Determining the Generator

The above is very nice, but how do you figure out what the generator is, explicitly? We have not found an explicit form for the transformation operators, so we cannot just Taylor expand it to find the generator. The answer is that we must Taylor expand the transformed wavefunction.

Let's begin by Taylor-expanding the transformation operator:

$$\begin{aligned}\psi_{q'}(q' = u) &= \langle q' = u \,|\psi\rangle = \langle q = u \,|\, T_P^\dagger |\psi\rangle \\ &\approx \langle q = u \,|\left(I + \frac{i}{\hbar}\,\varepsilon\, G\right)|\psi\rangle \\ &= \langle q = u \,|\psi\rangle + \frac{i}{\hbar}\,\varepsilon\,\langle q = u \,|G|\psi\rangle\end{aligned}$$

where we have done the usual substitution trick to avoid confusion and where the $\approx$ sign is used because we have considered an infinitesimal transformation by parameter value ε.

Generators for Continuous Coordinate Transformations (cont.)

Let's move things around:

$$\frac{i}{\hbar}\,\varepsilon\,\langle q = u\,|G|\psi\,\rangle \approx \langle q' = u\,|\psi\,\rangle - \langle q = u\,|\psi\,\rangle = \left[\langle q' = u\,| - \langle q = u\,|\right]|\psi\,\rangle$$

We know that, for any transformation, there is an equality relationship between the $\{|q'\,\rangle\}$ basis and the $\{|q\,\rangle\}$ basis that will allow us to rewrite $|q' = u\,\rangle$ in terms of $\{|q\,\rangle\}$ basis elements. Let's write this as $|q' = u\,\rangle = |q = q(q' = u, \varepsilon)\,\rangle$ where ε indicates the value of the transformation parameter that determines the relationship between q and q'. So we have

$$\begin{aligned}\frac{i}{\hbar}\,\varepsilon\,\langle q = u\,|G|\psi\,\rangle &\approx \left[\langle q = q(q' = u)\,| - \langle q = u\,|\right]|\psi\,\rangle \\ &= \psi_q(q = q(q' = u, \varepsilon)) - \psi_q(q = u)\end{aligned}$$

That is, we have now written everything on the right side in terms of the $\{|q\,\rangle\}$-basis wavefunction.

Generators for Continuous Coordinate Transformations (cont.)

Finally, because we are considering an infinitesimal transformation, we may calculate the right side by Taylor expansion in ε near $\varepsilon = 0$. Note that $\varepsilon = 0$ corresponds to $q = q'$, so we will be able to evaluate the derivatives at $q = q' = u$ and $\varepsilon = 0$. We have

$$\frac{i}{\hbar}\,\varepsilon\,\langle q = u\,|\,G|\psi\,\rangle \approx \varepsilon\,\left.\frac{\partial \psi_q}{\partial q}\right|_{q=u}\,\left.\frac{\partial q}{\partial \varepsilon}\right|_{\substack{q=q'=u \\ \varepsilon=0}}$$

where the derivatives in the last expression are schematic — there may be more than one coordinate, so one has to generalize appropriately to multiple coordinates. Regardless, one see that one can obtain the action of G on an arbitrary state, projected onto the $\{|q\,\rangle\}$ basis, which should let us determine explicitly what G is. We'll see this work in examples.

Example 12.3: Generator for the Passive Rotation Transformation

Recall from Example 12.2 that the position-basis wavefunctions of the untransformed state in the $\{|q\rangle\}$ and $\{|q'\rangle\}$ bases are:

$$\psi_{ab}(x,y) = \langle x, y\,|\psi_{ab}\rangle = \sqrt{\frac{4}{L_1 L_2}}\,\sin\left(\frac{a\,\pi\,x}{L_1}\right)\,\sin\left(\frac{b\,\pi\,y}{L_2}\right)$$

$$\psi_{ab,q'}(x',y') = \langle x', y'\,|\psi_{ab}\rangle = \sqrt{\frac{4}{L_1 L_2}}\,\sin\left(\frac{a\,\pi\,(x'c_\theta - y's_\theta)}{L_1}\right)\,\sin\left(\frac{b\,\pi\,(x's_\theta + y'c_\theta)}{L_2}\right)$$

Let's calculate the above Taylor expansion. First, let's recall that

$$x(x',y') = x'c_\theta - y's_\theta \qquad y(x',y') = x's_\theta + y'c_\theta$$

Taking the derivatives and allowing θ to be infinitesimal gives

$$\left.\frac{\partial x}{\partial\theta}\right|_{\substack{x=x'=u\\ y=y'=v\\ \theta=0}} \approx -v \qquad \left.\frac{\partial y}{\partial\theta}\right|_{\substack{x=x'=u\\ y=y'=v\\ \theta=0}} \approx u$$

So, our expression becomes

$$\frac{i}{\hbar}\,\theta\,\langle x=u, y=v\,|G|\psi\,\rangle = \theta\,\left.\frac{\partial\psi_{ab,q}}{\partial x}\right|_{\substack{x=u\\y=v}}\left.\frac{\partial x}{\partial\theta}\right|_{\substack{x=x',=u\\y=y',=v\\\theta=0}} + \theta\,\left.\frac{\partial\psi_{ab,q}}{\partial y}\right|_{\substack{x=u\\y=v}}\left.\frac{\partial y}{\partial\theta}\right|_{\substack{x=x',=u\\y=y',=v\\\theta=0}}$$

$$= \theta\left(-y\frac{\partial}{\partial x}+x\frac{\partial}{\partial y}\right)\psi_{ab,q}(x,y)\Big|_{\substack{x=u\\y=v}}$$

$$= \frac{i}{\hbar}\,\theta\langle x=u, y=v\,|\,(X\,P_y - Y\,P_x)\,|\psi_{ab}\,\rangle$$

where we used

$$i\,\hbar\,\frac{\partial}{\partial x}\psi_{ab,q}(x,y) = \langle x,y\,|P_x|\psi_{ab}\,\rangle \qquad i\,\hbar\,\frac{\partial}{\partial y}\psi_{ab,q}(x,y) = \langle x,y\,|P_y|\psi_{ab}\,\rangle$$

So we have $G = X\,P_y - Y\,P_x$. This expression should be familiar: straightforward replacement of classical x, y, p_x, p_y in $l_z = x\,p_y - y\,p_x$ with the corresponding quantum mechanical operators yields $L_z = X\,P_y - Y\,P_x$. Hence, $G = L_z$, the z-axis angular momentum operator, and we see that L_z generates rotations about the z axis, $T(\theta) = \exp\left(-\frac{i}{\hbar}\,\theta\,L_z\right)$.

Active Coordinate Transformations

We have defined passive coordinate transformations as a simple relabeling of space. However, one could have viewed the transformation as a movement of all the particles and potentials in the problem relative to the original axes. This is called an active transformation. In contrast to the passive transformation case, we transform the states and the Hamiltonian also. We call the new state $|\psi'\rangle = T_A|\psi\rangle$ where T_A is the operator that maps from the old states to the new states (the active transformation operator).

That said, we will need a way to define explicitly what we mean by the above. The natural way to do that is to consider a set of transformed axes $\{q'\}$ with corresponding new position and momentum basis elements $\{|q'\rangle\}$ and $\{|p'_q\rangle\}$, and to define the transformation in terms of its action on the basis elements in the same way as we did for passive transformations,

$$e^{i\theta}|q'\rangle = T_A|q\rangle$$

where θ is again real, and the prefactor exists for the same reasons discussed in connection to passive transformations. We shall take $\theta = 0$ in general. Though the definition has the same form, there is indeed a distinction between passive and active transformation operators that we will explain shortly.

Active Coordinate Transformations (cont.)

The unitarity of T_A follows by an argument similar to that used for T_P. Since we are just moving the system relative to the underlying space, the dimensionality of the new basis must be the same as that of the old basis; that is, the mapping is one-to-one, or invertible. Assuming the new basis elements are normalized in the same way as the old ones, the transformation also preserves inner products between basis elements because there are no coefficients in front. Therefore, it preserves all inner products, and hence is unitary, $T_A^\dagger = T_A^{-1}$.

The unitarity of T_A then lets us see in a more intuitive way how the state transforms. The transformed state satisfies

$$\langle q' \,|\psi'\rangle = \langle q \,|\, T_A^\dagger T_A \,|\psi\rangle = \langle q \,|\, T_A^{-1} T_A \,|\psi\rangle = \langle q \,|\psi\rangle$$

That is, the projection of the transformed state onto the transformed basis elements is the same as the projection of the untransformed state onto the untransformed basis elements. Since this projection is the position-basis wavefunction, what we are saying is that the position-basis wavefunction for the transformed state depends on the new (primed) coordinate axes in the same way as the position-basis wavefunction for the untransformed state depended on the old (unprimed) coordinate axes.

Active Coordinate Transformations (cont.)

The definition of transformed operators follows in the same way as it did for the passive transformation: because we have new coordinate axes and new (position- and momentum-) basis elements, we want to **define** new operators that act on the transformed basis elements in the same way as the old operators acted on the untransformed basis elements:

$$\langle q_1' \,|Q'|q_2' \rangle \equiv \langle q_1 \,|Q\,|q_2 \rangle \tag{12.2}$$

By the definition of the transformation operator's action on basis elements, $|q' \rangle = T_A|q \rangle$, we also have $\langle q_1' \,|Q'|q_2' \rangle = \langle q_1 \,|T_A^\dagger \, Q' \, T_A|q_2 \rangle$. Combining the two statements gives

$$\langle q_1 \,|Q\,|q_2 \rangle = \langle q_1 \,|T_A^\dagger \, Q' \, T_A \,|q_2 \rangle$$

Since this relation holds for all $|q_1 \rangle$ and $|q_2 \rangle$, it therefore holds that

$$Q' = T_A \, Q \, T_A^\dagger = T_A \, Q \, T_P^{-1}$$

The above proof carries through for any operator, including the $\{P_q\}$, and thus we now have a means to transform the operators. Notice how the formula is identical to the one we calculated for the passive transformation case; the distinction between active and passive transformations will be explained below. Again, if there is any confusion, one only needs to write the above in terms of matrix elements.

Active Coordinate Transformations (cont.)

In contrast to the passive transformation case, we *will* transform H to H' because we want to move the particles and potentials. Unitarity assures us that the transformed Hamiltonian's eigenstates are the transformed eigenstates of the untransformed Hamiltonian. That is, if $|\psi_E\rangle$ is an eigenstate of H with eigenvalue E, then we have

$$H'\left(T_A|\psi_E\rangle\right) = \left(T_A H\, T_A^\dagger\right)\left(T_A|\psi_E\rangle\right) = T_A H\,|\psi_E\rangle = T_A E\,|\psi_E\rangle = E\left(T_A|\psi_E\rangle\right)$$

For continuous transformations, we may write the transformation operator in terms of generators. Recall that we were able to write the action of the transformation operator on the basis elements and operators as

$$|q'\rangle = \exp\left(-\frac{i}{\hbar}\,\varepsilon\, G\right)|q\rangle \qquad O' = \exp\left(-\frac{i}{\hbar}\,\varepsilon\, G\right) O \exp\left(\frac{i}{\hbar}\,\varepsilon\, G\right)$$

The above forms continue to hold for active transformations, though, as we will see, one may not assume that the sign of ε will be the same for passive and active transformations. In addition, we may now also write

$$|\psi'\rangle = \exp\left(-\frac{i}{\hbar}\,\varepsilon\, G\right)|\psi\rangle \qquad H' = \exp\left(-\frac{i}{\hbar}\,\varepsilon\, G\right) H \exp\left(\frac{i}{\hbar}\,\varepsilon\, G\right)$$

Active Coordinate Transformations (cont.)

The next obvious question to ask is: are the untransformed eigenstates of the untransformed Hamiltonian, the $\{|\psi_E\rangle\}$, also eigenstates of the transformed Hamiltonian, and are the eigenstates of the transformed Hamiltonian, the $\{|\psi'_E\rangle = T_A|\psi_E\rangle\}$, also eigenstates of the untransformed Hamiltonian? The answer to both questions is, in general, no:

$$H\,|\psi'_E\rangle = H\,T_A|\psi_E\rangle \qquad H'|\psi_E\rangle = H\,T_A^\dagger|\psi'_E\rangle \tag{12.3}$$

We see that we need $[H, T_A] = 0$ in order for the above to simplify in the necessary fashion for $|\psi'_E\rangle$ to be an eigenstate of H and for $|\psi_E\rangle$ to be an eigenstate of H'. We shall discuss later symmetry transformations, which do satisfy the above commutation property and for which the answer to the above questions is "yes".

Relation between Passive and Active Transformations

It is clear that passive and active transformations are very much alike. Why make any distinction at all? What is the distinction?

In the passive case, we define a new set of coordinate axes and corresponding basis elements, leave the states and Hamiltonian unchanged (*i.e.*, leave the particles and potentials fixed relative to the old coordinate axes), and ask what the position-basis wavefunctions looks like in terms of the new coordinate axes and what the Hamiltonian looks like in terms of the new operators. That is, we are interested in $\langle q' | \psi \rangle$ and $H\left(Q', P_q'\right)$.

In the active case, we define a new set of coordinate axes and corresponding basis elements and transform the states and Hamiltonians along with the basis elements, and then we *will* ask what the new position-basis wavefunction looks like in terms of the old coordinate axes and what the Hamiltonian looks like in terms of the old operators. That is, we will be interested in $\langle q | \psi' \rangle$ and $H'(Q, P_q)$.

Before writing any formulae, let's first think conceptually about what the difference is, with our example of a coordinate system rotation for a particle in a 2d box in mind, Example 12.2.

Active Coordinate Transformations (cont.)

For the passive transformation, we rotate the axes CCW by an angle θ and leave the box and state in place. We calculate the position-basis wavefunction in terms of the new coordinates and the Hamiltonian in terms of the new operators. The wavefunction and Hamiltonian look ugly because the box is at an angle $-\theta$ with respect to the new coordinates.

For the active transformation, we rotate the axes *and the box and state* CCW by an angle θ. We will see that the new wavefunction in terms of the old coordinates and the Hamiltonian in terms of the old operators are ugly because the box and wavefunction are rotated by an angle $+\theta$ relative to the old coordinate axes.

We see that there is a sign flip of the transformation parameter involved because the box is at an angle $-\theta$ relative to the new coordinate axes for the passive transformation while the box is at an angle $+\theta$ relative to the old coordinate axes for the active transformation. This is in general what happens, that the passive and active transformations are related by a sign flip of the transformation parameter. Effectively, $T_P = T_A^{-1}$ and $T_A = T_P^{-1}$.

For discrete transformations like the mirror transformation in Example 12.1, the distinction vanishes because doing the transformation twice returns one to the original situation, so $T_P^2 = I$ and $T_A^2 = I$. Since the transformation is its own inverse, $T_P^{-1} = T_P$ and $T_A^{-1} = T_A$ and thus our relation $T_P = T_A^{-1}$ yields $T_P = T_A$. In cases of discrete transformations where $T_P^2 \neq I$ or $T_A^2 \neq I$, there will be a distinction and one will require $T_P = T_A^{-1}$ in order for the two transformations to yield the same effects.

Active Coordinate Transformations (cont.)

Now, let's write generic formulae that state the above:

$$\begin{aligned}
\text{Passive:}\quad & \psi_{q'}(q') = \langle q' \,|\psi\rangle = \langle q\,|\,T_P^\dagger|\psi\rangle \\
& H = \mathcal{H}(Q, P_q) = \mathcal{H}\left(T_P^\dagger Q' T_P,\, T_P^\dagger P_q' T_P\right) \\
\text{Active:}\quad & \psi_q'(q) = \langle q\,|\psi'\rangle = \langle q\,|\,T_A|\psi\rangle \\
& H' = T_A\, H\, T_A^\dagger = \mathcal{H}(Q', P_q') = \mathcal{H}\left(T_A Q T_A^\dagger,\, T_A P_q T_A^\dagger\right)
\end{aligned}$$

where we use $\mathcal{H}$ to indicate the classical Hamiltonian function, now treated as a function whose arguments can be operators; introducing $\mathcal{H}$ is necessary in order to be completely clear about what we mean. We immediately see that the old wavefunction in the new basis, $\langle q' \,|\psi\rangle$, and the new wavefunction in the old basis, $\langle q\,|\psi'\rangle$, are the same function of their arguments if and only if $T_P^\dagger = T_A$, which is equivalent to $T_P^{-1} = T_A$, the condition we stated above. Similarly, the untransformed Hamiltonian is a function of the new operators in the same way that the transformed Hamiltonian is a function of the old operators if and only if the same condition is met, $T_P^\dagger = T_A$.

An important implication of the above is that, if one focuses on the case $T_P^\dagger = T_A$ for the above reason, then the transformed operators will not be the same in the two cases, but will only have the same form when $T_P = T_A$.

Active Coordinate Transformations (cont.)

For continuous transformations, which can be written as complex exponentials of generators, our explicit forms are:

$$T_P(\varepsilon) = \exp\left(-\frac{i}{\hbar}\,\varepsilon\, G\right) \qquad T_A(\varepsilon) = \exp\left(\frac{i}{\hbar}\,\varepsilon\, G\right)$$

where we define the sign of ε be positive when the new axes are "to positive ε" of the old axes. For rotations, positive ε thus corresponds to the new axes being CCW from the old ones, as one would expect. For translation, positive ε corresponds to the origin displaced to a positive value in the particular coordinate for which the translation is being considered.

For such continuous transformations, the expressions given on the previous page for the transformation of the wavefunction can be simplified if the $\{|g\rangle\}$ basis is used:

$$\text{Passive:} \quad \psi_{g'}(g') = \langle g'|\psi\rangle = \langle g|\, e^{\frac{i}{\hbar}\,\varepsilon\, G}\,|\psi\rangle = e^{\frac{i}{\hbar}\,\varepsilon\, g}\,\langle g|\psi\rangle = e^{\frac{i}{\hbar}\,\varepsilon\, g}\,\psi_g(g)$$

$$\text{Active:} \quad \psi'_g(g) = \langle g|\psi'\rangle = \langle g|\, e^{\frac{i}{\hbar}\,\varepsilon\, G}\,|\psi\rangle = e^{\frac{i}{\hbar}\,\varepsilon\, g}\,\langle g|\psi\rangle = e^{\frac{i}{\hbar}\,\varepsilon\, g}\,\psi_g(g)$$

Lecture 38:
Coordinate Transformations:
Examples of Passive vs. Active Coordinate Transformations
Symmetry Transformations

Date Revised: 2009/01/30

Date Given: 2009/01/28

Example 12.4: Active Mirror Transformation of a Particle in a Two-Dimensional Box

Let's consider the active-transformation version of Example 12.1. The active version corresponds to inverting the box through the line $x = y$, putting the L_1 dimension along the y axis and the L_2 dimension along x.

We define a new coordinate system by $x' = y$ and $y' = x$ in the same way as for the passive transformation. We have the transformed basis $\{|x', y'\rangle = T_A\,|x, y\rangle\}$ (we drop the unity-modulus factor again), transformed operators $\{Q' = T_A\,Q\,T_A^\dagger\}$ and $\{P_q' = T_A\,P_q\,T_A^\dagger\}$, and the transformed eigenstate $|\psi_{ab}'\rangle = T_A\,|\psi_{ab}\rangle$. As we noted earlier in our general discussion, the action of T_A on basis elements and operators is identical to that of the corresponding T_P operator.

Active Coordinate Transformations (cont.)

Let's study the various wavefunctions we get out of these transformations. Remember first that, because of the unitarity of T_A, the projection of the transformed state onto the transformed basis is equal to the projection of the untransformed state onto the corresponding untransformed basis element:

$$\langle x' = u, y' = v \,|\psi'_{ab}\rangle = \langle x = u, y = v \,|\, T_A^\dagger \, T_A \,|\psi_{ab}\rangle = \langle x = u, y = v \,|\psi_{ab}\rangle$$

Note that the basis elements on the left and right side are related by $|x' = u, y' = v\rangle = T_A|x = u, y = v\rangle$, not by an equality! Another way of saying this is that we are not projecting onto $|x = u, y = v\rangle = |x' = v, y' = u\rangle$ on the left side of the equation.

Let's go through all the same things we did for the passive transformation example, starting with the wavefunction. The transformed basis elements $\{|x', y'\rangle\}$ are no different from those we found in the passive transformation case, so we still have

$$\psi_{ab,q}(x, y) = \sqrt{\frac{4}{L_1 L_2}} \sin\left(\frac{a\,\pi\,x}{L_1}\right) \sin\left(\frac{b\,\pi\,y}{L_2}\right)$$

$$\psi_{ab,q'}(x', y') = \sqrt{\frac{4}{L_1 L_2}} \sin\left(\frac{b\,\pi\,x'}{L_2}\right) \sin\left(\frac{a\,\pi\,y'}{L_1}\right)$$

Active Coordinate Transformations (cont.)

The new twist is that we have the new state $|\psi'_{ab}\rangle$ that is defined by $\langle x' = u, y' = v\,|\psi'_{ab}\rangle = \langle x = u, y = v\,|\psi_{ab}\rangle$. We may immediately write

$$\begin{aligned}
\psi'_{ab,q'}(x' = u, y' = v) &= \psi_{ab}(x = u, y = v) \\
&= \sqrt{\frac{4}{L_1 L_2}}\,\sin\left(\frac{a\,\pi\,u}{L_1}\right)\,\sin\left(\frac{b\,\pi\,v}{L_2}\right) \\
&= \sqrt{\frac{4}{L_1 L_2}}\,\sin\left(\frac{a\,\pi\,x'}{L_1}\right)\,\sin\left(\frac{b\,\pi\,y'}{L_2}\right) \\
\text{or,}\quad \psi'_{ab,q'}(x', y') &= \sqrt{\frac{4}{L_1 L_2}}\,\sin\left(\frac{a\,\pi\,x'}{L_1}\right)\,\sin\left(\frac{b\,\pi\,y'}{L_2}\right)
\end{aligned}$$

The functional dependence of $\psi'_{ab,q'}$ on (x', y') is the same as that of $\psi_{ab,q}$ on (x, y), as we have stated in our general discussion. Note of course that $\psi'_{ab,q'}(x', y') \neq \psi_{ab,q'}(x', y')$ because $|\psi'\rangle$ is a different state than $|\psi\rangle$ and hence their projections onto the same $\{|x', y'\rangle\}$ basis are different.

Active Coordinate Transformations (cont.)

Let's look at the last thing to consider, $\psi'_{ab,q} = \langle x = u, y = v \,|\psi'\rangle$:

$$\begin{aligned}\langle x = u, y = v \,|\psi'_{ab}\rangle &= \langle x' = v, y' = u \,|\psi'_{ab}\rangle \\ &= \sqrt{\frac{4}{L_1 L_2}} \sin\left(\frac{a\pi v}{L_1}\right) \sin\left(\frac{b\pi u}{L_2}\right) \\ \Longrightarrow \quad \psi'_{ab,q}(x,y) &= \sqrt{\frac{4}{L_1 L_2}} \sin\left(\frac{b\pi x}{L_2}\right) \sin\left(\frac{a\pi y}{L_1}\right)\end{aligned}$$

Thus, we see a relationship between the active and passive transformations:

$$\text{Passive:} \quad \psi_{ab,q'}(x',y') = \sqrt{\frac{4}{L_1 L_2}} \sin\left(\frac{b\pi x'}{L_2}\right) \sin\left(\frac{a\pi y'}{L_1}\right)$$

$$\text{Active:} \quad \psi'_{ab,q}(x,y) = \sqrt{\frac{4}{L_1 L_2}} \sin\left(\frac{b\pi x}{L_2}\right) \sin\left(\frac{a\pi y}{L_1}\right)$$

We thus have that the functional dependence of the transformed-basis wavefunction of the untransformed state $\psi_{ab,q'}$ on the transformed coordinates is the same as that of the untransformed-basis wavefunction on the untransformed coordinates $\psi'_{ab,q}$. This is a surprising and unusual result, though one for which there is a good reason.

Active Coordinate Transformations (cont.)

The above wavefunctions are the two projections $\langle x = u, y = v \,|\psi'_{ab}\,\rangle$ and $\langle x' = u, y' = v \,|\psi_{ab}\,\rangle$. (We again use the dummy variables (u, v) to avoid confusion about what arguments we need to consider.) In general, they are not related to each other because they may be written as

$$\langle x = u, y = v \,|\psi'_{ab}\,\rangle = \langle x = u, y = v \,|\, T_A |\psi_{ab}\,\rangle$$
$$\langle x' = u, y' = v \,|\psi_{ab}\,\rangle = \langle x = u, y = v \,|\, T_A^\dagger |\psi_{ab}\,\rangle = \langle x = u, y = v \,|\, T_P^\dagger |\psi_{ab}\,\rangle$$

However, in this specific case of a mirror transformation, they are related because $T_A^2 = I$ and thus $T_A^\dagger = T_A = T_P = T_P^\dagger$. As we noted in our general discussion, the generic result will be that the above kind of relation holds when the passive and active transformations are related by $T_P = T_A^\dagger = T_A^{-1}$.

Active Coordinate Transformations (cont.)

Recall that the passively and actively transformed operators will in general be equal when $T_P = T_A$, but in this case they will also be equal when $T_P = T_A^{-1}$ because $T_A^2 = I$. That is, we have

$$T_P O T_P^\dagger = T_A O T_A^\dagger \quad \text{in general}$$

$$T_P O T_P^\dagger = T_A^{-1} O \left(T_A^{-1}\right)^\dagger = T_A^\dagger O T_A \quad \text{for this particular case}$$

In particular, the relation between transformed and untransformed operators will not depend on whether we consider $T_P = T_A$ or $T_P = T_A^{-1}$: we always get $X' = Y$, $Y' = X$, $P_x' = P_y$ and $P_y' = P_x$.

Active Coordinate Transformations (cont.)

Let's now consider the Hamiltonian operator. We have from before

$$H(Q', P'_q) = \frac{[P'_x]^2 + [P'_y]^2}{2\,m} + V(Y', X')$$

For the active transformation, we have

$$H'(Q, P_q) = \frac{[P_x]^2 + [P_y]^2}{2\,m} + V(Y, X)$$

(Note: when we write H as a function of Q' and P'_q, we intend "take $H(Q, P_q)$ and use the relationships between Q, P_q and Q', P'_q to substitute for Q and P'_q;" we do not mean "replace Q, P_q in H with Q', P'_q directly." The same kind of statement holds for writing H' as a function of Q and P_q. However, when we write V, we do mean to treat it as a simple function of its arguments.) We see that the dependence of the untransformed Hamiltonian and the transformed operators is the same as that of the transformed Hamiltonian on the untransformed operators. This is again a special case resulting from the fact $T_P^2 = I$ and $T_A^2 = I$. In general, the above functional dependences will not be the same and one will have to require instead $T_P = T_A^{-1}$ in order for them to match up.

Active Coordinate Transformations (cont.)

One last comment on eigenstates. We know from our general discussion that $|\psi'_{ab}\rangle = T_A|\psi_{ab}\rangle$ is an eigenstate of H' with energy $E_{ab} = \hbar^2\pi^2\left(a^2/L_1^2 + b^2/L_2^2\right)/2\,m$. It is fairly obvious that $|\psi_{ab}\rangle$ is not an eigenstate of H' and $|\psi'_{ab}\rangle$ is not an eigenstate of H because the functional dependences of the corresponding wavefunctions on the coordinates do not match up with the potential in the corresponding Hamiltonian. So, though $\psi_{ab,q'}(x', y')$ may depend on its arguments in the same way that $\psi'_{ab,q}(x, y)$ depends on its arguments, these are not somehow the same state. They are different states simply because $|\psi_{ab}\rangle \neq |\psi'_{ab}\rangle$ no matter what basis one projects them on to, and of course one can only test equality between two states by projecting them onto the *same* basis. We can see from the work above that $\langle x, y\,|\psi'_{ab}\rangle \neq \langle x, y\,|\psi_{ab}\rangle$ and $\langle x', y'\,|\psi'_{ab}\rangle \neq \langle x', y'\,|\psi_{ab}\rangle$.

Example 12.5: Active Rotation Transformation of a Particle in a Two-Dimensional Box

Now, let's go through the same rigamarole for the active rotation transformation of the particle in a box. We'll go pretty quickly since we have a lot of experience in doing these transformations by now.

Let's write down the four possible wavefunctions:

$$\psi_{ab,q}(x,y) = \langle x, y \,|\psi_{ab}\rangle = \sqrt{\frac{4}{L_1 L_2}} \sin\left(\frac{a\,\pi\,x}{L_1}\right) \sin\left(\frac{b\,\pi\,y}{L_2}\right)$$

$$\psi_{ab,q'}(x',y') = \langle x', y' \,|\psi_{ab}\rangle = \sqrt{\frac{4}{L_1 L_2}} \sin\left(\frac{a\,\pi\,(x'c_\theta - y's_\theta)}{L_1}\right) \sin\left(\frac{b\,\pi\,(x's_\theta + y'c_\theta)}{L_2}\right)$$

$$\psi'_{ab,q}(x,y) = \langle x, y \,|\psi'_{ab}\rangle = \sqrt{\frac{4}{L_1 L_2}} \sin\left(\frac{a\,\pi\,(x\,c_\theta + y\,s_\theta)}{L_1}\right) \sin\left(\frac{b\,\pi\,(-x\,s_\theta + y\,c_\theta)}{L_2}\right)$$

$$\psi'_{ab,q'}(x',y') = \langle x', y' \,|\psi'_{ab}\rangle = \sqrt{\frac{4}{L_1 L_2}} \sin\left(\frac{a\,\pi\,x'}{L_1}\right) \sin\left(\frac{b\,\pi\,y'}{L_2}\right)$$

Active Coordinate Transformations (cont.)

Let's explicitly derive the third one because it is the only interesting one we have not dealt with before:

$$\begin{aligned}
\psi'_{ab,q}(x=u, y=v) &= \langle x=u, y=v \,|\psi'_{ab}\rangle \\
&= \langle x' = u\,c_\theta + v\,s_\theta, y' = -u\,s_\theta + v\,c_\theta \,|\psi'_{ab}\rangle \\
&= \psi'_{ab,q'}(x' = u\,c_\theta + v\,s_\theta, y' = -u\,s_\theta + v\,c_\theta) \\
&= \sqrt{\frac{4}{L_1 L_2}} \sin\left(\frac{a\,\pi\,(u\,c_\theta + v\,s_\theta)}{L_1}\right) \sin\left(\frac{b\,\pi\,(-u\,s_\theta + v\,c_\theta)}{L_2}\right)
\end{aligned}$$

where the equality in the first line between untransformed and transformed basis elements is the same equality we justified in Example 12.2.

Going back to the four wavefunctions, we clearly see that the dependence of the untransformed wavefunction on the transformed coordinates, $\psi_{ab,q'}(x', y')$, is the same as that of the transformed wavefunction on the untransformed coordinates, $\psi'_{ab,q}(x, y)$, if and only if we take the angle for the latter active transformation to be the opposite of the angle for the former passive transformation. That is, as we said for the general case, we require $T_P = T_A^{-1}$ in order to get these functional dependences to match up.

Active Coordinate Transformations (cont.)

As we noted earlier, when we take $T_P = T_A$, the transformed operators are the same regardless of whether the transformation is active of passive. We remind you that these relations are

$$X' = X\,c_\theta + Y\,s_\theta \qquad P_x' = P_x\,c_\theta + P_y\,s_\theta$$
$$Y' = -X\,s_\theta + Y\,c_\theta \qquad P_y' = -P_x\,s_\theta + P_y\,c_\theta$$

Of course, if we want the wavefunction functional dependences to match up, then the passively and actively transformed operators will **not** be equal because they will have different signs on θ.

Active Coordinate Transformations (cont.)

Let's look at the relation between the untransformed and transformed Hamiltonians:

$$H(Q,P_q) = \frac{[P_x]^2 + [P_y]^2}{2\,m} + V(X,Y)$$

$$H(Q',P'_q) = \frac{[P'_x]^2 + [P'_y]^2}{2\,m} + V\left(X'c_\theta - Y's_\theta, X's_\theta + Y'c_\theta\right)$$

$$H'(Q,P_q) = \frac{[P_x]^2 + [P_y]^2}{2\,m} + V\left(X\,c_\theta + Y\,s_\theta, -X\,s_\theta + Y\,c_\theta\right)$$

$$H(Q',P'_q) = \frac{[P'_x]^2 + [P'_y]^2}{2\,m} + V\left(X',Y'\right)$$

We have obtained $H'(Q',P'_q)$ by $H' = T_A H T_A^\dagger$, which we calculate by inserting $T_A^\dagger T_A$ between every power of untransformed operators and thus transforming them. The result should be intuitively clear, though, because it just consists of replacing every operator O in H with its transformed version O'. We see that the functional dependence of H on X' and Y' only matches up with the functional dependence of H' on X and Y when the sign of θ for the active and passive trasnformations are opposite, as we expect from our general discussion.

Passive and Active Transformations Use the Same Operator

Note that we now no longer need to distinguish between active and passive transformations at an operator level, as we saw that they act on the bases in the same way. The distinction between passive and active transformations is more a matter of which things one looks at: in the passive case, one cares about writing the untransformed state in terms of the transformed basis elements and the untransformed Hamiltonian in terms of the transformed operators; in the active case, one focused on writing the transformed state in terms of the untransformed basis elements and the transformed Hamiltonian in terms of the untransformed operators. When one compares the two, one sees that the passive and active transformations corresponding to a given transformation operator T do different things. But this is a matter of interpretation, not a matter of a distinction between passive and active transformations at an operator level.

Lecture 39:
Coordinate Transformations:
Symmetry Transformations

Date Revised: 2009/01/30
Date Given: 2009/01/30

Definition

A coordinate transformation T is a symmetry transformation of a Hamiltonian H if it satisfies any of the three following equivalent statements

$$H' = T\,H\,T^\dagger = H \qquad \Longleftrightarrow \qquad T\,H = H\,T \qquad \Longleftrightarrow \qquad [T,H] = 0 \qquad (12.4)$$

That is, a symmetry transformation is a coordinate transformation that commutes with the Hamiltonian. In practice, one must write H and H' in terms of the same operators in order to demonstrate the equality. That is, given

$$H(Q,P_q) = \mathcal{H}(Q,P_q) \qquad H'(Q',P_q') = \mathcal{H}(Q',P_q')$$
$$H(Q',P_q') = \mathcal{H}(Q = T^\dagger Q' T, P_q = T^\dagger P_q' T)$$
$$H'(Q,P_q) = \mathcal{H}(Q' = TQT^\dagger, P_q' = T\,P_q\,T^\dagger)$$

where $\mathcal{H}(q,p_q)$ is the classical Hamiltonian function, one needs to demonstrate one of the equalities

$$H'(Q,P_q) = H(Q,P_q) \qquad H(Q',P_q') = H'(Q',P_q')$$

This is the active transformation method for showing a coordinate transformation is a symmetry transformation.

Symmetry Transformations (cont.)

There is also a passive transformation method. If and only if the above equalities hold, we then have

$$H(Q', P_q') = H'(Q', P_q') = \mathcal{H}(Q', P_q')$$

We also have $H(Q, P_q) = \mathcal{H}(Q, P_q)$. So, if H depends on Q', P_q' in the same way as it depends on Q, P_q, then T is a symmetry transformation. **This is the passive transformation method for showing a coordinate transformation is a symmetry transformation.**

Since each set of eigenstates forms a complete basis, it must therefore hold that the eigenstates of H and H' are the same. More generally, we can conclude

$$\langle \psi | H' | \psi \rangle = \langle \psi | H | \psi \rangle = \langle \psi' | H | \psi' \rangle$$

Moreover, eigenstates of H must be eigenstates of T and vice versa — the two operators are simultaneously diagonalizable. Not only are eigenstates of H unaffected by the symmetry transformation up to a unity-modulus factor (the eigenvalue, which must be unity modulus for unitary operators), but the eigenstates of T are therefore also eigenstates of H. Therefore, if a system is put in an eigenstate of T at $t = 0$, it remains in that state for all time.

Symmetry Transformations (cont.)

The above fact is a useful classification tool for eigenstates of H. It also sometimes allows us to show that specific matrix elements of some operator between eigenstates of H vanish without explicit calculation due to its properties under the T symmetry transformation. When we get to perturbation theory later, this will be useful for showing that some contributing perturbing matrix elements vanish.

We must remember one thing, though, about how far symmetry transformations go. Though we have stated above that eigenstates of H and H' are the same, we shall see in examples that this does not necessarily mean that the position-basis wavefunctions of the eigenstates will be unchanged by the transformation.

Symmetry Transformations (cont.)

More explicitly: because of unitarity, we always have that the functional dependence on the transformed coordinates of the position-basis wavefunction in the transformed basis of a transformed eigenstate is the same as that of the position-basis wavefunction in the untransformed basis of an untransformed eigenstate on the untransformed coordinates,

$$\psi'_{E,q'}(q' = u) = \langle q' = u \,|\psi'_E\rangle = \langle q = u \,|\psi_E\rangle = \psi_{E,q}(q = u)$$

However, it does not in general hold that the position-basis wavefunction in the *transformed basis* of an *untransformed eigenstate* depends on the *transformed coordinates* in the same way as the untransformed eigenstate's position-basis wavefunction in the untransformed basis depends on the untransformed coordinates:

$$\psi_{E,q'}(q' = u) = \langle q' = u \,|\psi_E\rangle \neq \langle q = u \,|\psi_E\rangle = \psi_{E,q}(q = u) \quad \text{in general}$$

Similarly, it does not in general hold that the position-basis wavefunction in the *untransformed basis* of a *transformed eigenstate* depends on the *untransformed coordinates* in the same way as the transformed eigenstate's position-basis wavefunction in the transformed basis depends on the transformed coordinates:

$$\psi'_{E,q}(q = u) = \langle q = u \,|\psi'_E\rangle \neq \langle q' = u \,|\psi'_E\rangle = \psi'_{E,q'}(q' = u) \quad \text{in general}$$

So, we expect states to change their position-basis representations under symmetry transformations, but for their energy eigenvalues to stay the same.

Symmetry Transformations (cont.)

Continuous Symmetry Transformations

Suppose a continuous coordinate transformation $T(\varepsilon)$ is a symmetry transformation of a Hamiltonian H. As we have shown before, one way of writing this is $[T(\varepsilon), H] = 0$. Then, by considering the infinitesimal form of $T(\varepsilon)$, we conclude $[G, H] = 0$. This has a number of important implications:

- The eigenstates of G are eigenstates of H and vice versa — the two operators are simultaneously diagonalizable.
- Not only does $T(\varepsilon)$ commutes with the unitary time evolution operator, but so does G: $U = e^{\frac{i}{\hbar} H t}$, $[G, U(t)] = 0$.
- If the system is in an eigenstate $|g\rangle$ of G at $t = 0$, it remains in that eigenstate for all time.
- The expectation value of G in any state is conserved:

$$\begin{aligned}\langle\psi(t)|G|\psi(t)\rangle &= \langle\psi(0)|U^\dagger(t)\,GU(t)|\psi(0)\rangle \\ &= \langle\psi(0)|U^\dagger(t)\,U(t)\,G|\psi(0)\rangle = \langle\psi(0)|G|\psi(0)\rangle\end{aligned}$$

That is, in any way one can think of, the physical quantity associated with G is a conserved quantity. And, as with commutativity of T and H, this provides ways of classifying eigenstates of H and showing matrix elements of operators will vanish.

Example 12.6: Mirror Symmetry Transformation for a Particle in a Square Box

Consider our particle in a two-dimensional box, but now let the two sides of the box be the same size, $L_1 = L_2 \equiv L$. Let's first consider how H changes under the transformations. The symmetry of the potential ensures that $V(x, y) = V(y, x)$. We have for the passively and actively transformed Hamiltonians

$$H = \frac{[P_x']^2 + [P_y']^2}{2\,m} + V(Y', X') = \frac{[P_x']^2 + [P_y']^2}{2\,m} + V(X', Y')$$

$$H' = T\left(\frac{P_x^2 + P_y^2}{2\,m} + V(X, Y)\right) T^\dagger = \frac{(P_x')^2 + (P_y')^2}{2\,m} + V(X', Y')$$

$$= \frac{P_y^2 + P_x^2}{2\,m} + V(Y', X') = \frac{P_x^2 + P_y^2}{2\,m} + V(X, Y) = H$$

where in the second step for H we used $V(x, y) = V(y, x)$ and in the third step for H' we used $P_x' = P_y$, $P_y' = P_x$, and the symmetry of V and in the fourth step we used $X' = Y$ and $Y' = X$. The passive transformation method is the demonstration that H depends on X', Y', P_x', P_y' in the same way that it depends on X, Y, P_x, P_y. The equality of H' and H when both are written as a function of X, Y, P_x, P_y specifically demonstrates $T\,H\,T^\dagger = H$: this is the active transformation method. Both of these equalities arise because now the potential $V(x, y)$ is unchanged under the mirror transformation.

Symmetry Transformations (cont.)

Let's also look at the wavefunctions of the eigenstates, which are:

$$\psi_{ab,q}(x,y) = \frac{2}{L}\sin\left(\frac{a\pi x}{L}\right)\sin\left(\frac{b\pi y}{L}\right)$$

$$\psi_{ab,q'}(x',y') = \frac{2}{L}\sin\left(\frac{b\pi x'}{L}\right)\sin\left(\frac{a\pi y'}{L}\right)$$

$$\psi'_{ab,q'}(x',y') = \frac{2}{L}\sin\left(\frac{a\pi x'}{L}\right)\sin\left(\frac{b\pi y'}{L}\right)$$

$$\psi'_{ab,q}(x,y) = \frac{2}{L}\sin\left(\frac{b\pi x}{L}\right)\sin\left(\frac{a\pi y}{L}\right)$$

We see the point we noted above: the wavefunctions change under a symmetry transformation, but they retain the same energy eigenvalue. For example, $\psi_{ab,q'}(x',y')$ depends on its arguments in a different way than $\psi_{ab,q}(x,y)$. However, they are both position-basis projections of the same state $|\psi_{ab}\rangle$, which is an eigenstate of H and H' because the two are equal. Similarly, $\psi'_{ab,q}(x,y)$ depends on its arguments in a different way that $\psi'_{ab,q'}(x',y')$, but again they are both position-basis projections of the same state $|\psi'_{ab}\rangle$, which is an eigenstate of both H and H' because they are equal.

Example 12.7: Rotational Symmetry Transformation for a Particle in a Circular Box

Consider the classical potential

$$V(x,y) = \begin{cases} 0 & \sqrt{x^2+y^2} < R \\ \infty & \sqrt{x^2+y^2} \geq R \end{cases} \iff V(\rho) = \begin{cases} 0 & \rho < R \\ \infty & \rho \geq R \end{cases}$$

Let's first consider the effect of a rotation transformation by an angle θ CCW on the Hamiltonian.

Symmetry Transformations (cont.)

This is pretty trivial: we know that, while the rotation transformation yields the relations

$$X' = X\,c_\theta + Y\,s_\theta \qquad Y' - X\,s_\theta + Y\,c_\theta$$
$$P_x' = P_x c_\theta + P_y\,s_\theta \qquad P_y' = -P_x\,s_\theta + P_y c_\theta$$

it nevertheless holds that

$$\left[P_x'\right]^2 + \left[P_y'\right]^2 = [P_x]^2 + [P_y]^2 \qquad \left[X'\right]^2 + \left[Y'\right]^2 = [X]^2 + [Y]^2$$

Since $\mathcal{H}$ only depends on $p_x^2 + p_y^2$ and $x^2 + y^2$, it will therefore hold that

$$\text{Active:} \quad H'(X, Y, P_x, P_y) = H(X, Y, P_x, P_y)$$
$$\text{Passive:} \quad H(X', Y', P_x', P_y') = \mathcal{H}(X', Y', P_x', P_y')$$

both of which show that the rotation is a symmetry transformation of this Hamiltonian.

Symmetry Transformations (cont.)

Next, let's look at the effect of the rotation transformation on the wavefunctions. We must first know what the form of the wavefunction is. The eigenvalue-eigenvector equation in polar coordinates is

$$\left[-\frac{\hbar^2}{2\,m}\left(\frac{\partial^2}{\partial\rho^2}+\frac{1}{\rho}\,\frac{\partial}{\partial\rho}+\frac{1}{\rho^2}\,\frac{\partial^2}{\partial\phi^2}\right)+V(\rho)\right]\psi_q(\rho,\phi)=E\,\psi_q(\rho,\phi)$$

The eigenvalue-eigenvector problem is separable when written in terms of polar coordinates, $\psi_q(\rho,\phi)=P(\rho)\Phi(\phi)$:

$$\frac{d^2}{d\rho^2}\,P(\rho)+\frac{1}{\rho}\,\frac{d}{d\rho}\,P(\rho)+\frac{2\,m\,(E-V(\rho))}{\hbar^2}\,P(\rho)-\frac{m^2}{\rho^2}\,P(\rho)=0$$

$$\frac{d^2}{d\phi^2}\,\Phi(\phi)+m^2\Phi(\phi)=0$$

There will thus be solutions of the form $\psi_{Em,q}(\rho,\phi)=P_{E,m}(\rho)\Phi_m(\phi)$.

Symmetry Transformations (cont.)

The ϕ differential equation is trivially solved, giving

$$\Phi_m(\phi) = \frac{1}{\sqrt{2\,\pi}}\, e^{i\,m\,\phi} \qquad m \text{ any integer}$$

where the quantization on m arises because of the boundary condition $\Phi(\phi + 2\,\pi) = \Phi(\phi)$, which is required for $\Phi(\phi)$ to be a single-valued function of ϕ. It will not be necessary to solve the radial equation for our purposes here, so we will just leave the radial solutions in the form $P_{E,m}(\rho)$ where E designates the unspecified quantized energies (quantized by the boundary conditions on the wavefunction at $\rho = 0$ and $\rho = R$) and m is the ϕ quantum number, which sets the boundary condition on $P(\rho)$ at $\rho = 0$. Our eigenstates are thus of the form

$$\psi_{Em,q}(\rho, \phi) = \frac{1}{\sqrt{2\,\pi}}\, e^{i\,m\,\phi} P_{E,m}(\rho)$$

Symmetry Transformations (cont.)

Now, let's consider the effect of a rotation transformation. Since we are working in polar coordinates, the rotation transformation takes the form

$$\rho' = \rho \qquad \phi' = \phi - \theta$$

which corresponds to the x' axis being an angle θ CCW of the x axis, as we did in Examples 12.2 and 12.5. This yields the following relations between basis elements:

$$\begin{aligned} |\rho' = u, \phi' = v\rangle &= T\,|\rho = u, \phi = v\rangle \\ |\rho' = u, \phi' = v\rangle &= |\rho = u, \phi = v + \theta\rangle \\ |\rho' = u, \phi' = v - \theta\rangle &= |\rho = u, \phi = v\rangle \end{aligned}$$

Symmetry Transformations (cont.)

The various wavefunctions we usually consider will be

$$\psi_{nm,q}(\rho,\phi) = \frac{1}{\sqrt{2\,\pi}}\, e^{i\,m\,\phi} P_{E,m}(\rho)$$

$$\psi_{Em,q\,'}(\rho',\phi') = \frac{1}{\sqrt{2\,\pi}}\, e^{i\,n\,(\phi'+\theta)} P_{E,m}(\rho') = e^{i\,n\,\theta}\, \frac{1}{\sqrt{2\,\pi}}\, e^{i\,m\,\phi'} P_{E,m}(\rho')$$

$$\psi'_{Em,q\,'}(\rho',\phi') = \frac{1}{\sqrt{2\,\pi}}\, e^{i\,m\,\phi'} P_{E,m}(\rho')$$

$$\psi'_{Em,q}(\rho,\phi) = \frac{1}{\sqrt{2\,\pi}}\, e^{i\,n\,(\phi-\theta)} P_{E,m}(\rho) = e^{-i\,n\,\theta}\, \frac{1}{\sqrt{2\,\pi}}\, e^{i\,m\,\phi} P_{E,m}(\rho)$$

If it is unclear how we obtained the above, you may repeat the steps we did in the various prior examples using the above relations between the position-basis elements in the two bases. From the above, we thus see that a rotation transformation adds a unity-modulus factor $e^{\pm i\,n\,\theta}$ in front of the wavefunction. That factor of course can be absorbed in the normalization, so in fact we see that the functional dependence of $\psi_{Em,q\,'}(\rho',\phi')$ on ρ',ϕ' is the same as that of $\psi_{Em,q}(\rho,\phi)$ on ρ,ϕ, which is unusual and does not always happen to the wavefunction due to a symmetry transformation. The same holds in the comparison of $\psi'_{Em,q\,'}(\rho',\phi')$ to $\psi'_{Em,q}(\rho,\phi)$.

Lecture 40:
Coordinate Transformations:
Time Transformations

Date Revised: 2009/02/02
Date Given: 2009/02/02

Time Translation vs. Coordinate Transformations

At a fundamental level, the time coordinate is different from the spatial coordinates in non-relativistic classical and quantum mechanics. The main reason is that time is considered a parameter: in classical mechanics, it is not a dynamical variable; in QM, there is no observable operator associated with time. There is no momentum (operator) conjugate to time with which to construct a Poisson bracket (commutation) relation. Energy is, to some extent, the observable conjugate to time, but it is only a rough correspondence between (t, H) and (X, P).

Another problematic issue is that the Schrödinger Equation tells us how time evolution — *i.e.*, time translation — should be done. We don't have any freedom in defining the time translation. Or, we could assume we could define time translations as we like, but they would have no physical relevance. For example, it would be nonsense to define a transformation

$$T(t)\,|q(t=0)\,\rangle = |q'\,\rangle = |q(t)\,\rangle$$

because we know that, even for an explicitly time-independent Hamiltonian, $|q\,\rangle$ is in general not an eigenstate of H and so its time evolution is not a simple time-dependent mapping to other position-basis elements in the space.

Time Transformations (cont.)

That said, we can see that much of our formalism for coordinate transformations can be carried over to time translations. Let's define time translation to be the standard unitary time-evolution operator, but allowing for the fact that H may be time-dependent and may not commute with itself at different times ($[H(t), H(t')] = 0$ is only guaranteed for $t = t'$). That is,

$$T(t) = U(t) = \mathcal{T}\left[\exp\left(-\frac{i}{\hbar}\int_0^t dt'\, H(t')\right)\right] = \lim_{N\to\infty}\prod_{j=0}^{N-1}\exp\left[-\frac{i}{\hbar}\left(\frac{t}{N}\right)H\left(\frac{j}{N}\,t\right)\right]$$

(Equation 4.19) That is, there is a very specific transformation that provides time translation: the only freedom is in the amount of time t. The form of the translation depends on the particular Hamiltonian, unlike any of the coordinate transformations, which can be applied to any system with any Hamiltonian. We know the above transformation is unitary from previous work. We will write the transformation as

$$|\psi(t)\rangle = |\psi'\rangle = U(t)\,|\psi(t=0)\rangle$$

Time Transformations (cont.)

For an arbitrary operator, we take as a requirement

$$\langle \psi_1(t) \,|O(t)\,| \psi_2(t) \rangle \equiv \langle \psi_1' \,|O'|\psi_2' \rangle = \langle \psi_1(t=0) \,|O_0\,| \psi_2(t=0) \rangle$$
$$\Longleftrightarrow \qquad O(t) \equiv O' = U(t)\, O_0\, U^\dagger(t)$$

where we define our untransformed operator to be O_0 to avoid confusion between O_0 and $O(t)$ (*i.e.*, if we had taken our untransformed operator to be just O, as we did for coordinate transformations, it would be unclear whether O refers to $O(t)$ or $O(t=0)$). Of course, $O_0 = O(t=0)$. In addition to the above relation between the matrix elements of $O(t)$ in the transformed basis and the matrix elements of O_0 in the untransformed basis, we might be inclined to ask whether there is a relation between the matrix elements of $O(t)$ in the untransformed basis and the matrix elements of O_0 in the transformed basis. As we frequently find, there is no such relation in general:

$$\langle \psi_1(t) \,|O_0|\psi_2(t) \rangle = \langle \psi_1(t=0) \,|U^\dagger(t)\, O_0\, U(t)|\psi_2(t=0) \rangle$$
$$\langle \psi_1(t=0) \,|O(t)|\psi_2(t=0) \rangle = \langle \psi_1(t=0) \,|U(t)\, O_0\, U^\dagger(t)|\psi_2(t=0) \rangle$$

The expressions simplify and are equal if $[O_0, U(t)] = 0$ for all t, which is equivalent to $[O_0, H(t)] = 0$ for all t. But, in that case, $O(t) = O_0$ and O is conserved.

Time Transformations (cont.)

Even for a time-dependent Hamiltonian, it holds that the infinitesimal time evolution operator that goes from t to $t + dt$ is

$$U(t + dt, t) = 1 - \frac{i}{\hbar} H(t)\, dt$$

If H is time-dependent, knowing that H is the generator is not particuarly helpful because there is no fixed basis of eigenstates of H. So, at this point, let us specialize to time-independent H. The time translation operator has an explicit form,

$$U(t) = e^{-\frac{i}{\hbar} H t}$$

As note above, the usual coordinate transformation formula, $|q'\rangle = T|q\rangle$ does not hold because $T|q\rangle = e^{-\frac{i}{\hbar} H t}\,|q(t=0)\rangle$ is in general not a position basis element. The other two general relations do carry through:

$$|\psi'\rangle \equiv |\psi(t)\rangle = e^{-\frac{i}{\hbar} H t}\,|\psi(t=0)\rangle \equiv e^{-\frac{i}{\hbar} H t}\,|\psi\rangle$$

$$O(t) \equiv O' = e^{-\frac{i}{\hbar} H t}\, O\, e^{\frac{i}{\hbar} H t}$$

Time Transformations (cont.)

As we did for coordinate transformations, we are led to consider the effect of the transformation on the eigenstates of the generator:

$$|E(t)\rangle \equiv |E'\rangle = e^{-\frac{i}{\hbar}Ht}|E\rangle = e^{-\frac{i}{\hbar}Et}|E\rangle$$

We may also follow the example from coordinate transformations for transformation of the generator-eigenbasis representation of an arbitrary state $|\psi\rangle$. That is, consider the transformation of an arbitrary state $|\psi\rangle$ when written in the $\{|E\rangle\}$-basis representation. We denote this representation, $\langle E|\psi(t)\rangle \equiv \langle E|\psi'\rangle$, by $\psi_E(E,t) \equiv \psi'_E(E)$, and term it the "$\{|E\rangle\}$-basis wavefunction" or "energy-basis wavefunction" (it is **not** to be confused with our notation $|\psi_E\rangle$ for the Hamiltonian eigenstate with energy E). It is

$$\langle E|\psi(t)\rangle \equiv \langle E|\psi'\rangle = \langle E|e^{-\frac{i}{\hbar}Ht}|\psi\rangle = e^{-\frac{i}{\hbar}Et}\langle E|\psi\rangle = e^{-\frac{i}{\hbar}Et}\psi_E(E)$$

Time Transformations (cont.)

Consider the time transformation of matrix elements of the operators in the energy eigenbasis:

$$\langle E_1 \,|O(t)\,|E_2\,\rangle = \langle E_1 \,|O'|E_2\,\rangle = \langle E_1 \,|e^{-\frac{i}{\hbar}\,H\,t}\,O\,e^{\frac{i}{\hbar}\,H\,t}\,|E_2\,\rangle = e^{-\frac{i}{\hbar}(E_1-E_2)\,t}\langle E_1\,|O\,|E_2\,\rangle$$

So, in spite of the imperfect analogy between the time-translation transformation and coordinate transformations, we see that many of the general results for coordinate transformations carry through for time translation.

As a final note, we state the obvious: the generator H commutes with the Hamiltonian H, so the eigenstates of the Hamiltonian are eigenstates of the generator and vice versa, and the generator H is conserved in all the usual ways.

Time Transformations (cont.)

Time Reversal Transformation

The other analogy between time and coordinate transformations that we should consider is the temporal equivalent of a parity transformation, the time reversal transformation. As with translations, there are some subtleties that arise. One might think that the right rule for the transformation would be

$$\Pi_t |\psi(t)\rangle = |\psi(-t)\rangle$$

Let's see what time evolution equations $|\psi(-t)\rangle$ satisfies. To avoid confusion, define $|\phi(t)\rangle = |\psi(-t)\rangle$ and work on it, changing variables to $t' = -t$ and then back to t to make the manipulations clearer:

$$\begin{aligned} i\hbar \frac{d}{dt} |\phi(t)\rangle = i\hbar \frac{d}{dt} |\psi(-t)\rangle = -i\hbar \frac{d}{d(-t)} |\psi(-t)\rangle = -i\hbar \frac{d}{dt'} |\psi(t')\rangle \\ = -H(t') |\psi(t')\rangle = -H(-t) |\psi(-t)\rangle = -H(-t) |\phi(t)\rangle \end{aligned}$$

Time Transformations (cont.)

So, $|\phi(t)\rangle = |\psi(-t)\rangle$ satisfies the Schrödinger Equation with a Hamiltonian that is the negative of the time reverse of the Hamiltonian for which $|\psi(t)\rangle$ satisfies the Schrödinger Equation. This is clearly not useful if we consider an eigenstate of a time-independent H: if $|\psi(t)\rangle = |E(t)\rangle$ is an eigenstate of H with energy E, then $|\psi(-t)\rangle$ is an eigenstate of $-H$ with energy $-E$!

If we look at the Schrödinger Equation in the position basis, we are led to a slightly different definition of the operation of the parity operator. Suppose that $|\psi(t)\rangle$ is a solution of the Schrödinger Equation with Hamiltonian H. Then the position-basis Schrödinger Equation is

$$i\,\hbar\,\frac{d}{dt}\,\psi_x(x,t) = H\left(x, -i\,\hbar\,\frac{d}{dx}, t\right)\,\psi_x(x,t)$$

where we have shown H as a function of x and $-i\,\hbar\,\frac{d}{dx}$ instead of X and P because we are working in the position basis. Take the complex conjugate:

$$-i\,\hbar\,\frac{d}{dt}\,\psi_x^*(x,t) = H^*\left(x, -i\,\hbar\,\frac{d}{dx}, t\right)\,\psi_x^*(x,t)$$

Time Transformations (cont.)

Now, define $\phi_x(x,t) = \psi_x^*(x,-t)$. Then we have

$$-i\,\hbar\,\frac{d}{dt}\,\phi_x(x,-t) = H^*\left(x, -i\,\hbar\,\frac{d}{dx}, t\right)\,\phi_x(x,-t)$$

Change variables to $t' = -t$:

$$-i\,\hbar\,\frac{d}{d(-t')}\,\phi_x(x,t') = H^*\left(x, -i\,\hbar\,\frac{d}{dx}, -t'\right)\,\phi_x(x,t')$$

$$i\,\hbar\,\frac{d}{dt'}\,\phi_x(x,t') = H^*\left(x, -i\,\hbar\,\frac{d}{dx}, -t'\right)\,\phi_x(x,t')$$

So we see that $\phi_x(x,t)$ is the position-space representation of a state that satisfies the Schrödinger Equation for the Hamiltonian that is the time-reversed, complex conjugate of the original Hamiltonian. This is a much more reasonable result than what we had before.

Time Transformations (cont.)

So, we define

$$\Pi_t|\psi(t)\rangle = |\phi(t)\rangle \quad \text{with} \quad \phi_x(x,t) = \psi_x^*(x,-t) \Longleftrightarrow \langle x\,|\phi(t)\rangle = \langle \psi(-t)\,|x\rangle$$

The second part of the definition, $\langle x\,|\phi(t)\rangle = \langle \psi(-t)\,|x\rangle$, corrects what was wrong with the first definition, which would have defined $\langle x\,|\phi(t)\rangle = \langle x\,|\psi(-t)\rangle$. The subtlety here is that, while complex conjugation is perfectly well defined for the $\psi_x(x,-t)$, it is not defined for a Hilbert space vector $|\psi(-t)\rangle$, so we had to go through the position-basis representation of the Schrödinger Equation to see that the complex conjugation step was needed.

We see from the above that the condition for time reversal to be a symmetry of the Hamiltonian is for $H = H^*$; if this happens, then the Schrödinger Equation that $|\phi(t)\rangle$ satisfies in the position basis is the same as the one that $|\psi(t)\rangle$ satisfies.

It would be nice to be able to write this last statement (and, in general, transformation rules for operators under Π_t) in a form similar to what we did for the spatial parity transformation, but the lack of a direct ket-to-ket mapping makes this difficult.

Lecture 41:
Coordinate Transformations:
Relation between Classical and Quantum Transformations

Date Revised: 2009/02/02

Date Given: Not given; optional material

Canonical Transformations and Continuous Coordinate Transformations: General Relations

In classical mechanics, a canonical transformation is one that preserves Hamilton's equations. Consider a system described by a classical Hamiltonian $\mathcal{H}(q, p_q)$ and for which Hamilton's equations in q and p_q are satisfied

$$\frac{dp_q}{dt} = -\frac{\partial \mathcal{H}}{\partial q} \qquad \frac{dq}{dt} = \frac{\partial \mathcal{H}}{\partial p_q}$$

Consider the contact transformation defined by

$$q' = q'(q(t), p_q(t), t) \qquad p'_q = p'_q(q(t), p_q(t), t)$$

(It is called a contact transformation for historical reasons; simple coordinate transformations are a special case in which $q' = q'(q(t), t)$ and are called point transformations.)

The Relation between Classical and Quantum Transformations (cont.)

This contact transformation is a canonical transformation if, when the Hamiltonian is written in terms of the new coordinates and momenta, $\mathcal{H} = \mathcal{H}(q', p')$, Hamilton's equations are satisfied in the new coordinates,

$$\frac{dp'_q}{dt} = -\frac{\partial \mathcal{H}}{\partial q'} \qquad \frac{dq'}{dt} = \frac{\partial \mathcal{H}}{\partial p'_q}$$

One can show that a transformation is canonical if the Poisson brackets of the new coordinates in terms of the old coordinates is equal to 1:

$$\{q', p'_q\}_{q,p_q} = \frac{\partial q'}{\partial q}\frac{\partial p'_q}{\partial p_q} - \frac{\partial p'_q}{\partial q}\frac{\partial q'}{\partial p_q} = 1$$

The Relation between Classical and Quantum Transformations (cont.)

The quantum analogue is Postulate 2, whose alternate version (see Section 6.5) is

$$[Q, P_q] = i\,\hbar$$

If we carry the Poisson bracket over to a commutator in the standard fashion, the classical requirement becomes

$$[Q', P'_q] = i\,\hbar$$

We do not need to explicitly write a $_{Q,P_q}$ subscript because the only way to evaluate the commutator is to rewrite it in terms of Q and P_q. We thus see that the requirement that a classical contact transformation be canonical is exactly analogous to the requirement that Postulate 2 be satisfied by the transformed Q and P_q operators after a quantum mechanical coordinate transformation. We know that the this requirement is satsfied by any unitary transformation:

$$\begin{aligned}[Q', P'] &= Q'P' - P'Q' = T\,Q\,T^\dagger\,T\,P_q\,T^\dagger - T\,P_q\,T^\dagger\,T\,Q\,T^\dagger \\ &= T\,Q\,P_q\,T^\dagger - T\,P_q Q\,T^\dagger = T\,[Q, P_q]\,T^\dagger = T\,i\,\hbar\,T^\dagger = i\,\hbar\end{aligned}$$

Infinitesimal Classical and Quantum Transformations

An infinitesimal canonical transformation is of the form

$$q' = q + \delta\varepsilon \frac{\partial G}{\partial p_q} = q + \delta\varepsilon \{q, G\}_{q,p_q} \qquad p_q' = p_q - \delta\varepsilon \frac{\partial G}{\partial q} = p_q + \delta\varepsilon \{p_q, G\}_{q,p_q}$$

where $G = G(q, p_q)$ is some differentiable function of q and p_q satisfying $\frac{\partial^2 G}{\partial q\, \partial p_q} = \frac{\partial^2 G}{\partial p_q\, \partial q}$. The use of the same function G in the two equations ensures the transformation is canonical, as we can see by calculating the Poisson bracket:

$$\begin{aligned}
\{q', p_q'\}_{q,p_q} &= \frac{\partial q'}{\partial q}\frac{\partial p_q'}{\partial p_q} - \frac{\partial p_q'}{\partial q}\frac{\partial q'}{\partial p_q} \\
&= \left(1 + \delta\varepsilon \frac{\partial^2 G}{\partial q\, \partial p_q}\right)\left(1 - \delta\varepsilon \frac{\partial^2 G}{\partial p_q\, \partial q}\right) - \left(-\delta\varepsilon \frac{\partial^2 G}{\partial q^2}\right)\left(\delta\varepsilon \frac{\partial^2 G}{\partial p^2}\right) \\
&= 1 + \delta\varepsilon \left(\frac{\partial^2 G}{\partial q\, \partial p_q} - \frac{\partial^2 G}{\partial p_q\, \partial q}\right) + \delta\varepsilon^2 \left(-\frac{\partial^2 G}{\partial q\, \partial p_q}\frac{\partial^2 G}{\partial p_q\, \partial q} + \frac{\partial^2 G}{\partial q^2}\frac{\partial^2 G}{\partial p_q^2}\right)
\end{aligned}$$

The Relation between Classical and Quantum Transformations (cont.)

With our condition on $G(q, p_q)$, the coefficient of $\delta\varepsilon$ vanishes. We neglect the term of order $(\delta\varepsilon)^2$, leaving

$$\{q',\, p_q'\}_{q,p_q} = 1$$

The infinitesimal transformation may be rewritten as

$$\delta q = q' - q = \delta\varepsilon\,\{q, G\}_{q,p_q} \qquad \delta p_q = p_q' - p_q = \delta\varepsilon\,\{p, G\}_{q,p_q}$$

Now, compare this to the transformation of an operator Q or P_q by the infinitesimal coordinate transformation generated by the quantum analogue of G, $G(Q, P_q)$. The operator is

$$T(\delta\varepsilon) = I + \frac{i}{\hbar}\,\delta\varepsilon\, G$$

We have changed the sign on $\delta\varepsilon$ here for reasons that are more easily explained in terms of particular examples, which we shall do below.

The Relation between Classical and Quantum Transformations (cont.)

The infinitesimal transformation of an arbitrary operator O is

$$O' = T(\delta\varepsilon)\, O\, T^{\dagger}(\delta\varepsilon) = \left(I + \frac{i}{\hbar}\,\delta\varepsilon\, G\right) O \left(I - \frac{i}{\hbar}\,\delta\varepsilon\, G\right) = O + \frac{\delta\varepsilon}{i\,\hbar}\,[O, G] + \mathcal{O}(\delta\varepsilon^2)$$

or $\quad \delta O = O' - O = \dfrac{\delta\varepsilon}{i\,\hbar}\,[O, G]$

So, $\quad \delta Q = Q' - Q = \dfrac{\delta\varepsilon}{i\,\hbar}\,[Q, G] \qquad \delta P = P_q' - P_q = \dfrac{\delta\varepsilon}{i\,\hbar}\,[P_q, G]$

We see there is a direct correspondence: the rules for coordinate transformations in quantum mechanics directly correspond to the analogous rules in classical mechanics.

Here are a couple examples:

- **Spatial translation**
  The generator is $G(q, p_q) = p_q$ or $G(Q, P_q) = P_q$. The classical and quantum transformations are

$$\delta q = \delta\varepsilon \{q, p_q\}_{q,p_q} = \delta\varepsilon \qquad \delta p = \delta\varepsilon \{p_q, p_q\}_{q,p_q} = 0$$

$$\delta Q = \frac{\delta\varepsilon}{i\hbar}[Q, P_q] = \delta\varepsilon \qquad \delta P_q = \frac{\delta\varepsilon}{i\hbar}[P_q, P_q] = 0$$

  Now let us discuss the issue of the sign of $\delta\varepsilon$. First, consider the classical coordinate transformation. If $\delta\varepsilon > 0$, then we have increased the value of the coordinate q by $\delta\varepsilon$. This corresponds to a passive transformation in which we move the origin to $-q$ by a distance $\delta\varepsilon$, or to an active transformation in which we move the particle to $+q$ by a distance $\delta\varepsilon$.

The Relation between Classical and Quantum Transformations (cont.)

Now consider the quantum analogues. When we talked about a passive translation by $\delta\varepsilon$, we defined that to be such that the origin moved by $\delta\varepsilon$ to $+q$, resulting in $Q' = Q - \delta\varepsilon\, I$. For an active transformation, the the same result holds, and it makes sense as follows. If we consider $\langle\psi\,|Q'|\psi\,\rangle$, we are considering the position of the untransformed state relative to the new axes. Thus, we expect $\langle\psi\,|Q'|\psi\,\rangle = \langle\psi\,|\,(Q - \delta\varepsilon\, I)\,|\psi\,\rangle$. If we consider the relation between the expectation value of Q or Q' for the transformed and untransformed states, then we have $\langle\psi'\,|Q|\psi'\,\rangle = \langle\psi\,|Q\,|\psi\,\rangle + \delta\varepsilon$. Hence, $\langle Q\rangle \longrightarrow \langle Q\rangle + \delta\varepsilon$ is completely consistent with $Q' = Q - \delta\varepsilon\, I$ as well as with the classical transformation $q \longrightarrow q + \delta\varepsilon$.

So, we see that we must employ the opposite sign in the transformation of the operators as we do in the classical transformation in order to depict the same physical situation.

The Relation between Classical and Quantum Transformations (cont.)

- **Time translation**
  The generator is $G(q, p_q) = \mathcal{H}(q, p_q)$ or $G(Q, P_q) = H(Q, P_q)$. The classical transformations are

$$\delta q = \delta\varepsilon \{q, \mathcal{H}\}_{q,p_q} \qquad \delta p = \delta\varepsilon \{p_q, \mathcal{H}\}_{q,p_q}$$

  If we take $\delta\varepsilon = \delta t$, we have

$$\frac{\delta q}{\delta t} = \delta\varepsilon \{q, \mathcal{H}\}_{q,p_q} \qquad \frac{\delta p}{\delta t} = \delta\varepsilon \{p_q, \mathcal{H}\}_{q,p_q}$$

  which are just Hamilton's equations. For the quantum case, we have

$$\delta Q = \frac{\delta\varepsilon}{i\,\hbar}[Q, H] \qquad \delta P_q = \frac{\delta\varepsilon}{i\,\hbar}[P_q, H]$$

  Again, let $\delta\varepsilon = \delta t$ and also take expectation values. We thus have

$$\left\langle \frac{\delta Q}{\delta t} \right\rangle = \frac{1}{i\,\hbar}\langle [Q, H] \rangle \qquad \left\langle \frac{\delta P_q}{\delta t} \right\rangle = \frac{1}{i\,\hbar}\langle [P_q, H] \rangle$$

  and we thus recover Ehrenfest's Theorem.

  The issue with the sign of $\delta\varepsilon$ is identical to that for spatial translations.

Finite Classical and Quantum Transformations

One can construct both classical and quantum transformations from the infinitesimal transformations involving Poisson brackets or commutators. Since we are more familiar with the quantum transformation, let's do that one first. The standard form is

$$O' = e^{\frac{i}{\hbar}\varepsilon G} O e^{-\frac{i}{\hbar}\varepsilon G}$$

We may rewrite this as follows:

$$O' = O + \frac{\varepsilon}{i\hbar}[O, G] + \frac{1}{2!}\left(\frac{\varepsilon}{i\hbar}\right)^2 [[O, G], G] + \cdots$$

(You can check that this is correct by simply expanding it.)

The Relation between Classical and Quantum Transformations (cont.)

We can see that classical transformations can be written in similar fashion. Our general relation for some classical function $O(q, p_q)$ is

$$\delta O = \delta\varepsilon\{O, G\}_{q,p_q} \qquad \Longleftrightarrow \qquad \frac{d}{d\varepsilon} O(q, p_q; \varepsilon) = \{O, G\}_{q,p_q}$$

This holds for any function $O(q, p_q; \varepsilon)$, so apply it to $dO/d\varepsilon$:

$$\frac{d^2}{d\varepsilon} O(q, p_q; \varepsilon) = \frac{d}{d\varepsilon}\frac{dO}{d\varepsilon} = \left\{\frac{dO}{d\varepsilon}, G\right\}_{q,p_q} = \left\{\{O, G\}_{q,p_q}, G\right\}_{q,p_q}$$

One can see that this formula extends to any order derivative, so we now know the full Taylor expansion for $O(q, p_q)$ as a function of ε:

$$O(q, p_q; \varepsilon) = O(q, p_q; \varepsilon = 0) + \varepsilon\{O, G\}_{q,p_q} + \frac{1}{2!}\varepsilon^2\{\{O, G\}_{q,p_q}, G\}_{q,p_q}$$

So, again, we have direct correspondence with the substitution $\{\ ,\} \to \frac{1}{i\hbar}[\ ,\ ]$.

Classical and Quantum Symmetry Transformations and Conserved Quantities

With the above correspondences in hand, it is immediately evident that the conditions for classical and quantum transformations to be symmetry transformations are also in correspondence. The quantum condition is that the generator of the transformation commute with H so that H is unchanged by the symmetry transformation (which follows from the formulae involving commutators on the previous page):

$$[H, G] = 0 \qquad \Longleftrightarrow \qquad H' = H$$

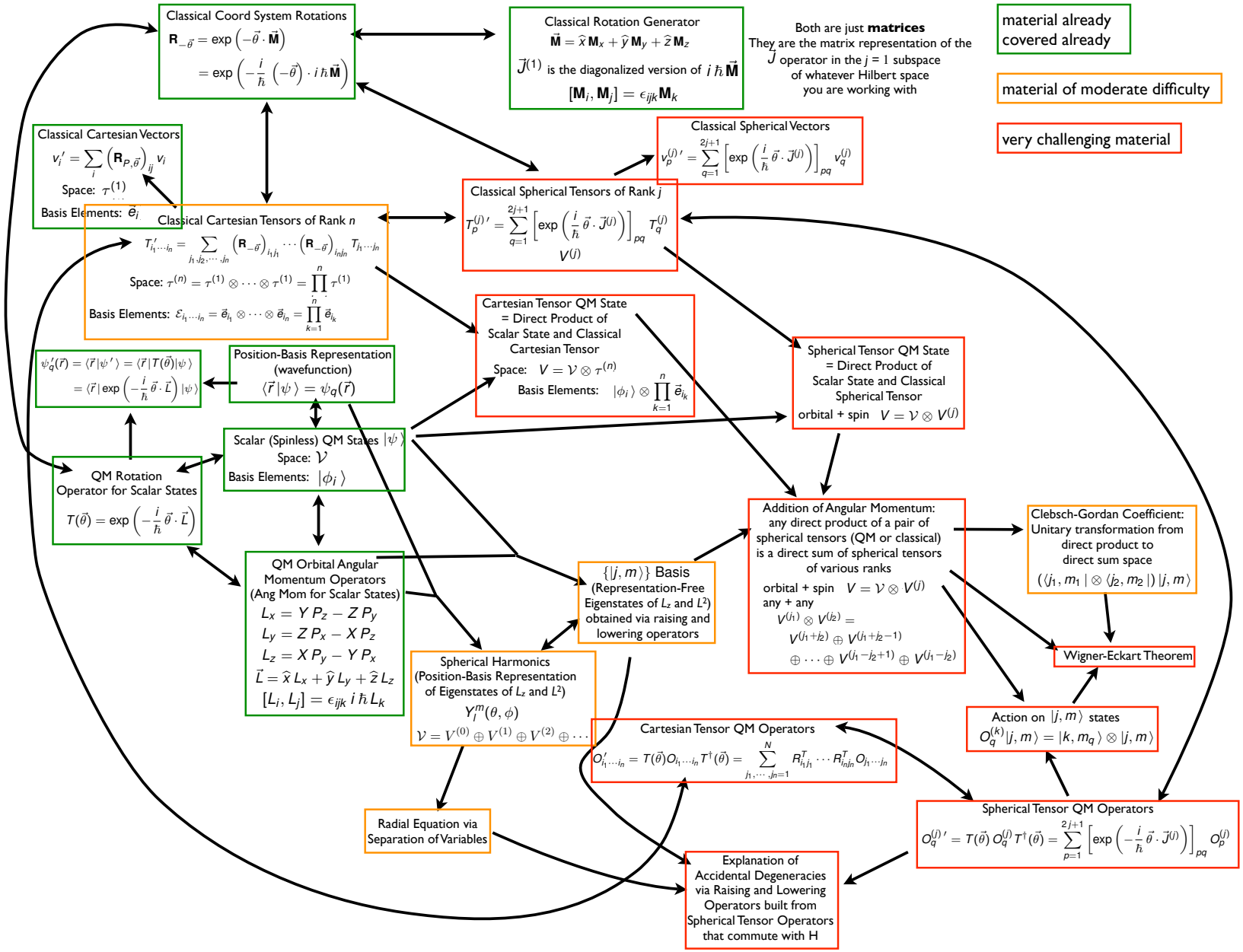
We see from the previous page that the classical Hamiltonian $\mathcal{H}$ is unchanged if a similar condition holds involving Poisson brackets:

$$\{H, G\}_{q,p_q} = 0 \qquad \Longleftrightarrow \qquad \mathcal{H}' = \mathcal{H}$$

Moreover, because $\mathcal{H}$ and H are the generators of time translation in the classical and quantum formalisms, respectively, the relation $[O, H] =$ or $\{O, \mathcal{H}\}_{q,p_q} = 0$ implies that O remains unchanged under time translation. Classically, the functional value O is conserved. Quantum mechanically, the operator O does not change with time and its matrix elements and expectation values are constant.

Section 13
Angular Momentum Summary

Angular Momentum Summary



Angular Momentum Summary (cont.)

What will make this material difficult:

- Many of the proofs and arguments are indirect. The crux of this whole business is the notion that the way classical and quantum mechanical states transform under rotations provides a useful way to classify them and allows you to do calculations in a relatively painless way. We thus focus on **transformation properties** of states rather than doing direct things like solving differential equations or calculating matrix elements explicitly, which is what we have spent most of the course doing.
- You will be forced to think conceptually because, as with transformations, notation only gets you so far. You **must** understand what the symbols you are writing down **mean**. This is the greatest problem I see students having — that they are not mastering the definitions and concepts, so they are unsure what certain symbols and notation mean.
- There will be a great deal of looking at different representations of the same objects by writing the Hilbert space in differing ways. An example is the idea of breaking down the Hilbert space of states of a particle in three spatial dimensions into the direct product of a space the describes the radial behavior and one that describes the angular behavior, and the further decomposition of the latter into subspaces of well-defined orbital angular momentum. Another examples is addition of angular momentum, wherein we take the direct product space of two angular momenta and break it down into a direct sum of the subspaces of well-defined total angular momentum.

Section 14
Rotations and Orbital Angular Momentum

Lecture 42:
Rotations and Orbital Angular Momentum in Two Dimensions

Date Revised: 2009/02/04
Date Given: 2009/02/04

Plan of Attack

We will study the problem of rotations and orbital angular momentum in the following sequence:

- **Rotation Transformations in Two Dimensions**
  We will first review classical rotation transformations in two dimensions, derive the formula for the active rotation transformation of a quantum mechanical state, and show that the generator of the transformation is the quantum analogue of the classical z-axis angular momentum, L_z.
- **The L_z Eigenvector-Eigenvalue Problem**
  L_z will be a Hermitian, observable operator. For Hamiltonians for which $[H, L_z] = 0$ – *i.e.*, Hamiltonians with rotational symmetry in two dimensions – H and L_z are simultaneously diagonalizable. Therefore, eigenvectors of H must also be eigenvectors of L_z, and so the eigenvectors of L_z will be of interest. We calculate the eigenvectors and eigenvalues of L_z and see how the requirement that eigenvectors of H be eigenvectors of L_z reduces the Schrödinger Equation to a differential equation in the radial coordinate only.
- **Rotation Transformations in Three Dimensions**
  We then generalize classical rotation transformations to three dimensions and use correspondences to identify the three angular momentum operators L_x, L_y, and L_z, as well as the total angular momentum magnitude L^2.

Plan of Attack (cont.)

- **The L^2-L_z Eigenvalue Problem**
  In three dimensions, we shall see that L_x, L_y, L_z, and L^2 are all Hermitian, observable operators. But no two of L_x, L_y, and L_z commute, while each of them commutes with L^2, so it becomes clear that useful set of operators to work with for Hamiltonians that are rotationally invariant in three dimensions is H, L_z, and L^2. We therefore consider the joint eigenvector-eigenvalue problem of L^2 and L_z and determine how it reduces the Schrödinger Equation to a differential equation in the radial coordinate only.

We will refer back frequently to material on continuous symmetry transformations that we covered in Section 12, so please review that material.

Passive Classical Rotation Transformations in Two Dimensions

A *passive* coordinate system rotation in two dimensions by an angle θ counterclockwise yields the following relationship between the components of a vector $\vec{a}$ in the untransformed system (a_x, a_y, a_z) and its components in the transformed system $(a_{x'}, a_{y'}, a_{z'})$:

$$a_{x'} = a_x\, c_\theta + a_y\, s_\theta \qquad a_{y'} = -a_x\, s_\theta + a_y\, c_\theta \qquad a_{z'} = a_z$$

where $c_\theta = \cos\theta$ and $s_\theta = \sin\theta$ as usual. The x' and y' axes are obtained by rotating the x and y axes counterclockwise by the angle θ. The rotation is termed passive because we are not changing the vector $\vec{a}$, we are simply writing its representation in terms of a new set of coordinate axes. The above may be written as a matrix operation:

$$\begin{bmatrix} a_{x'} \\ a_{y'} \\ a_{z'} \end{bmatrix} = \begin{bmatrix} c_\theta & s_\theta & 0 \\ -s_\theta & c_\theta & 0 \\ 0 & & 1 \end{bmatrix} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \equiv \mathbf{R}_{P,\theta\hat{z}} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}$$

where we use the $_P$ subscript to indicate a passive transformation (as we did in the QM case) and the $_{\theta\hat{z}}$ subscript to indicate the rotation angle from the untransformed to the transformed system.

Rotation Transformations in Two Dimensions (cont.)

Let us emphasize here the concept of **coordinate representations** of classical vectors. The unprimed and primed coordinate systems are just two different ways of labeling space. The vector $\vec{a}$ has not changed by relabeling space. However, the components of $\vec{a}$ in the two coordinate systems are different. We thus call (a_x, a_y, a_z) and $(a_{x'}, a_{y'}, a_{z'})$ two different **coordinate representations** of the same vector $\vec{a}$. This is very much the same idea as our discussion of different position-basis representations of a state $|\psi\rangle$ depending on whether we project it onto the position-basis elements for the original coordinate system $\{|x, y\rangle\}$ or those of the transformed coordinate system $\{x', y'\}$, giving position-basis representations $\langle x, y|\psi\rangle$ and $\langle x', y'|\psi\rangle$, respectively.

Active Classical Rotation Transformations in Two Dimensions

The classical analogue of an *active* coordinate transformation is to change the vector; that is, to fix the coordinate system and to change the vector by changing its coordinate representation (components) in that coordinate system. If we denote the **new vector** by $\vec{a}\,'$, then the coordinate representation (components) of $\vec{a}\,'$ are related to those of $\vec{a}$ by

$$\begin{bmatrix} a'_x \\ a'_y \\ a'_z \end{bmatrix} = \begin{bmatrix} c_\theta & -s_\theta & 0 \\ s_\theta & c_\theta & 0 \\ 0 & & 1 \end{bmatrix} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \equiv \mathbf{R}_{A,\theta\hat{z}} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}$$

or

$$a'_x = a_x\, c_\theta - a_y\, s_\theta \qquad a'_y = a_x\, s_\theta + a_y\, c_\theta \qquad a'_z = a_z$$

where both are being represented in the untransformed coordinate system. This transformation corresponds to physically rotating $\vec{a}$ by θ CCW about $\hat{z}$. $\vec{a}\,'$ is a **different** vector than $\vec{a}$ because its coordinate representation in this fixed coordinate system is different from that of $\vec{a}$. Again, this is in direct analogy to our active transformations in QM, where we kept the position basis unchanged but transformed the state, $|\psi\,'\rangle = T|\psi\rangle$, and saw that the states had different position-basis representations in the same basis, $\langle x, y\,|\psi\rangle$ and $\langle x, y\,|\psi\,'\rangle$.

Passive vs. Active Classical Rotation Transformations

The key difference between active and passive transformations is that the active transformation rotates the vector $\vec{a}$, creating a new vector $\vec{a}\,'$, while the passive transformation rotates the coordinate system so that the representation of the vector $\vec{a}$ changes from (a_x, a_y, a_z) to $(a_{x'}, a_{y'}, a_{z'})$, but the vector $\vec{a}$ is unchanged. This is in exactly analogy to what we considered for QM states: for a passive transformation, we consider the projection of the untransformed state $|\psi\rangle$ onto the transformed position basis $\{|q'\rangle = T|q\rangle\}$ by looking at $\langle q'|\psi\rangle$, while, for an active transformation, we consider the projection of the transformed state $|\psi'\rangle = T|\psi\rangle$ onto the untransformed basis $\{|q\rangle\}$ by looking at $\langle q|\psi'\rangle$.

It may be helpful to realize that the unit vectors of the transformed system, $\widehat{x}'$, $\widehat{y}'$, and $\widehat{z}'$, are obtained by performing an active transformation on the unit vectors of the untransformed system, $\widehat{x}$, $\widehat{y}$, and $\widehat{z}$.

Rotation Transformations in Two Dimensions (cont.)

The mathematical difference between the passive and active transformations is just the change of sign of the s_θ terms; that is $R_{P,-\theta\hat{z}} = R_{A,\theta\hat{z}}$. This sign flip tells us that the coordinate representation of $\vec{a}$ in a transformed coordinate system is literally equal to the coordinate representation in the untransformed coordinate system of the vector $\vec{a}'$ that has been obtained from $\vec{a}$ by active rotation by $-\theta\hat{z}$. Of course, in spite of this equality, we know $\vec{a}$ and $\vec{a}'$ are different vectors because the coordinate representations that are equal are coordinate representations in different coordinate systems (the transformed and untransformed systems). This is analogous to the situation in quantum mechanics of a passively transformed state having the same position-basis representation in the transformed basis as an actively transformed state has in the untransformed basis when the actively transformed state has been transformed using the inverse transformation as was used for the passive transformation (see Section 12.3).

It is convention to use $\mathbf{R}_{\theta\hat{z}}$ for $\mathbf{R}_{A,\theta\hat{z}}$ and to never use $\mathbf{R}_{P,\theta\hat{z}}$. We will follow this convention.

Generators for Classical Rotation Transformations in Two Dimensions

Since we are going to be considering generators in the quantum case and for the three-dimensional classical case, it is worth showing how the above transformation can be written as an operator exponential of a generator. As we did in connection with identifying the generator of a continuous coordinate transformation of quantum mechanical states, we will begin by considering an infinitesimal version of the above coordinate transformation:

$$\mathbf{R}_{\delta\theta\hat{z}} = \begin{bmatrix} \cos\delta\theta & -\sin\delta\theta & 0 \\ \sin\delta\theta & \cos\delta\theta & 0 \\ 0 & & 1 \end{bmatrix} \approx \begin{bmatrix} 1 & -\delta\theta & 0 \\ \delta\theta & 1 & 0 \\ 0 & & 1 \end{bmatrix}$$

Rotation Transformations in Two Dimensions (cont.)

The generic relationship between a classical coordinate transformation and its generator is

$$\mathbf{T}_\epsilon = \mathbf{I} + \epsilon\,\mathbf{G}$$

Instead of relating Hermitian generators to unitary coordinate transformation operators, we must relate antisymmetric generators to orthogonal coordinate transformation operators. (The generator must be antisymmetric, not symmetric, because we have no i in the argument of the exponential as we do for the QM version). Thus, it makes sense to rewrite our infinitesimal rotation operators as

$$\mathbf{R}_{\delta\theta\widehat{z}} = \mathbf{I} + \delta\theta\,\mathbf{M}_z \qquad \mathbf{M}_z \equiv \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, $\mathbf{M}_z$ is the classical generator of rotations about $\widehat{z}$. The use of the $_z$ subscript of course foreshadows similar operators for rotations about $\widehat{x}$ and $\widehat{y}$.

Rotation Transformations in Two Dimensions (cont.)

We of course recover the finite classical rotation transformation by the appropriate infinite product, yielding an exponential:

$$\mathbf{R}_{\theta\widehat{z}} = \lim_{N\to\infty}\left(\mathbf{I} + \frac{\theta}{N}\,\mathbf{M}_z\right)^N = \exp\left(\theta\,\mathbf{M}_z\right)$$

We may evaluate the above using the fact

$$\mathbf{M}_z^2 = -\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

This yields

$$\begin{aligned}
\mathbf{R}_{\theta\hat{z}} &= \sum_{n=0}^{\infty} \frac{\theta^n}{n!}\,\mathbf{M}_z^n = \mathbf{I} + \theta\,\mathbf{M}_z + \sum_{n=1}^{\infty}\left(\frac{\theta^{2n}}{(2n)!}\mathbf{M}_z^{2n} + \frac{\theta^{2n+1}}{(2n+1)!}\mathbf{M}_z^{2n}\mathbf{M}_z\right) \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \sum_{n=0}^{\infty} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \left(\frac{\theta^{2n}(-1)^n}{(2n)!} + \frac{\theta^{2n+1}(-1)^n}{(2n+1)!}\mathbf{M}_z\right) \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} (c_\theta + s_\theta\,\mathbf{M}_z) \\
&= \begin{bmatrix} c_\theta & -s_\theta & 0 \\ s_\theta & c_\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}$$

as expected.

Rotation Transformations in Two Dimensions (cont.)

What is the significance of the $\mathbf{M}_z$ matrix? See:

$$-\vec{r}^{\,T}\mathbf{M}_z\vec{p} = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix}$$
$$= x\,p_y - y\,p_x = \ell_z$$

That is, $\mathbf{M}_z$ can be used to compute the z component of the angular momentum when combined with the $\vec{r}$ and $\vec{p}$ vectors. $\mathbf{M}_z$ is in some nontrivial way connected to the z component of angular momentum.

Lecture 43:
Rotations and Orbital Angular Momentum in Two Dimensions, ct'd
The Eigenvalue Problem of L_z in Two Dimensions

Date Revised: 2009/02/06
Date Given: 2009/02/06

Quantum Mechanical Active Rotation Transformation in Two Dimensions

Let's recall Examples 12.2, 12.5, and 12.3 which were passive and active QM rotation and calculation of the generator of rotations. Recall that we found that the generator of the QM rotation transformation was the z-axis angular momentum operator,

$$G = X\,P_y - Y\,P_x \equiv L_z$$

and that the generic explicit form for the quantum mechanical operator for rotation transformations about the z-axis is

$$T(\theta\widehat{z}) = \exp\left(-\frac{i}{\hbar}\,\theta\,L_z\right)$$

We also now see the connection between the classical and quantum rotation formalisms. We saw that the $\mathbf{M}_z$ matrix that generates two-dimensional rotations returns the classical l_z when it acts on $\vec{r}$ and $\vec{p}$, $l_z = \vec{r}^{\,T}\mathbf{M}_z\vec{p}$. Thus, it is perhaps not surprising that the quantum generator of two-dimensional rotations is the quantum analogue, L_z.

Rotation Transformations in Two Dimensions (cont.)

To build some intuition about what exactly the above operator does, let's write down the projection of its action on a state onto the position basis and convert that to polar coordinates. We begin with the Cartesian coordinate version:

$$\frac{i}{\hbar}\,\theta\,\langle x,y\,|L_z|\psi\,\rangle = \frac{i}{\hbar}\,\theta\,\langle x,y\,|\,(X\,P_y - Y\,P_x)\,|\psi\,\rangle = \theta\left(x\,\frac{\partial}{\partial y} - y\,\frac{\partial}{\partial x}\right)\psi_q(x,y)$$

Now, we need to change variables to polar coordinates. The functional relationship between polar and cartesian coordinates is

$$\rho = \sqrt{x^2+y^2} \qquad \phi = \arctan\frac{y}{x}$$

Hence

$$\frac{\partial\rho}{\partial x} = \frac{x}{\rho} \qquad \frac{\partial\rho}{\partial y} = \frac{y}{\rho} \qquad \frac{\partial\phi}{\partial x} = \frac{-y/x^2}{1+(y/x)^2} = -\frac{y}{\rho^2} \qquad \frac{\partial\phi}{\partial y} = \frac{1/x}{1+(y/x)^2} = \frac{x}{\rho^2}$$

Rotation Transformations in Two Dimensions (cont.)

The chain rule thus tells us

$$\frac{\partial}{\partial x} = \frac{\partial \rho}{\partial x}\frac{\partial}{\partial \rho} + \frac{\partial \phi}{\partial x}\frac{\partial}{\partial \phi} = \frac{x}{\rho}\frac{\partial}{\partial \rho} - \frac{y}{\rho^2}\frac{\partial}{\partial \phi}$$
$$\frac{\partial}{\partial y} = \frac{\partial \rho}{\partial y}\frac{\partial}{\partial \rho} + \frac{\partial \phi}{\partial y}\frac{\partial}{\partial \phi} = \frac{y}{\rho}\frac{\partial}{\partial \rho} + \frac{x}{\rho^2}\frac{\partial}{\partial \phi}$$

So, then,

$$\theta\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\psi_q(x,y) = \theta\left(\frac{\partial}{\partial \phi}\right)\psi_q(\rho,\phi)$$

where we simply rewrite ψ_q in terms of ρ and ϕ using $x = \rho\cos\phi$ and $y = \rho\sin\phi$. So,

$$\begin{aligned}\langle x,y\,|\,\frac{i}{\hbar}\,\theta\,L_z|\psi\,\rangle &= \langle x,y\,|\,\frac{i}{\hbar}\,\theta\,(X\,P_y - Y\,P_x)\,|\psi\,\rangle \\ &= \langle x,y\,|\,\theta\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)|\psi\,\rangle = \langle x,y\,|\,\theta\,\frac{\partial}{\partial \phi}|\psi\,\rangle\end{aligned}$$

which looks like the action of the generator of a translation in the polar angle ϕ by an angle θ, as we expect.

Rotation Transformations in Two Dimensions (cont.)

Recall that we also calculated the action of the rotation transformation on the standard operators X, Y, P_x, and P_Y in Example 12.2 and 12.5. There we did it by calculating the matrix elements of the transformed operators in the untransformed basis, but that is tedious. We can do it much more quickly using operator arithmetic relations now that we know what the generator of the transformation is. We will need the relation

$$e^{-A}\, B\, e^{A} = B + [B, A] + \frac{1}{2!}\,[[B, A], A] + + \frac{1}{3!}\,[[[B, A], A], A] \cdots$$

(which we do not prove here). This relation will make use of the following important commutators:

$$\begin{aligned}
[X, L_z] &= [X, X\, P_y] - [X, Y\, P_x] = 0 - Y\,[X, P_x] = -i\,\hbar\, Y \\
[Y, L_z] &= [Y, X\, P_y] - [Y, Y\, P_x] = X\,[Y, P_y] - 0 = i\,\hbar X \\
[P_x, L_z] &= [P_x, X\, P_y] - [P_x, Y\, P_x] = [P_x, X]\, P_y - 0 = -i\,\hbar\, P_y \\
[P_y, L_z] &= [P_y, X\, P_y] - [P_y, Y\, P_x] = 0 - [P_y, Y]\, P_x = i\,\hbar\, P_x
\end{aligned}$$

The evident cyclicity of the above relations will be written succinctly when we consider rotations in three dimensions.

Rotation Transformations in Two Dimensions (cont.)

With the above, the transformed operators are easily evaluated:

$$
\begin{aligned}
X' &= T(\theta\hat{z})\, X\, T(-\theta\hat{z}) = e^{-\frac{i}{\hbar}\,\theta\, L_z}\, X\, e^{\frac{i}{\hbar}\,\theta\, L_z} \\
&= X + \frac{i}{\hbar}\,\theta[X, L_z] + \frac{1}{2!}\left(\frac{i}{\hbar}\,\theta\right)^2 [[X, L_z], L_z] + \frac{1}{3!}\left(\frac{i}{\hbar}\,\theta\right)^3 [[[X, L_z], L_z], L_z] + \cdots \\
&= X + \theta\, Y + \frac{\theta^2}{2!}\,(-1)\, X + \frac{\theta^3}{3!}\,(-1)\, Y + \cdots \\
&= X\left(1 - \frac{\theta^2}{2!} + \cdots\right) + Y\left(\theta - \frac{\theta^3}{3!} + \cdots\right) \\
&= X\, c_\theta + Y\, s_\theta \\
Y' &= -X\, s_\theta + Y\, c_\theta \\
P_x' &= P_x\, c_\theta + P_y\, s_\theta \\
P_y' &= -P_x\, s_\theta + P_y\, c_\theta
\end{aligned}
$$

where the last three are evaluated in the same way as X'.

Note the way in which X and Y are mixed and P_x and P_y are mixed to obtain the transformed operators; this looks very much like the transformation of classical vectors. We will discuss **vector** and **tensor** operators later.

Eigenvalues and Eigenfunctions of L_z

It is interesting to determine the eigenvalues and eigenfunctions of any Hermitian operator. It will be especially useful for L_z because of the many Hamiltonians with which it commutes.

We begin with the obvious, the eigenvector-eigenvalue equation for L_z:

$$L_z|\ell_z\rangle = \ell_z\,|\ell_z\rangle$$

where ℓ_z is an eigenvalue and $|\ell_z\rangle$ is the corresponding eigenvector. We take the product on the left with position basis elements since that is the basis in which we know the matrix elements of L_z from our previous calculation:

$$-i\,\hbar\left(x\,\frac{\partial}{\partial y} - y\,\frac{\partial}{\partial x}\right)\psi_{\ell_z}(x,y) = \ell_z\,\psi_{\ell_z}(x,y)$$

We also make use of the change of variables to (ρ,ϕ) to obtain a simpler equation:

$$-i\,\hbar\,\frac{\partial}{\partial\phi}\,\psi_{\ell_z}(\rho,\phi) = \ell_z\,\psi_{\ell_z}(\rho,\phi)$$

The Eigenvalue Problem of L_z in Two Dimensions (cont.)

The solution is obvious,

$$\psi_{\ell_z}(\rho,\phi) = R(\rho)\, e^{\frac{i}{\hbar}\,\phi\,\ell_z}$$

To this point, we have made no restrictions on ℓ_z. Not only are there no bounds or discretization, there is no prohibition against an imaginary component because ϕ is restricted to $[0, 2\,\pi]$ and so the exponential will not diverge. Hermiticity will obviously result in ℓ_z being real. Less obviously, it will also discretize ℓ_z. Recall from Section 3.9 that Hermiticity for the K (and hence P) operator for a particle on one dimension on a finite interval $[a, b]$ placed the requirement that any valid wavefunction vanish at the endpoints, $\psi(x = a) = 0 = \psi(x = b)$. We can derive a similar requirement here. Hermiticity implies

$$\langle \psi_1 \,|L_z|\psi_2 \,\rangle = \langle \psi_2 \,|L_z|\psi_1 \,\rangle^*$$

which, written out in terms of the position-basis wavefunction, is

$$\int_0^\infty d\rho\,\rho \int_0^{2\,\pi} d\phi\,\psi_1^*(\rho,\phi)\left(-i\,\hbar\,\frac{\partial}{\partial\phi}\right)\psi_2(\rho,\phi)$$
$$= \left[\int_0^\infty d\rho\,\rho \int_0^{2\,\pi} d\phi\,\psi_2^*(\rho,\phi)\left(-i\,\hbar\,\frac{\partial}{\partial\phi}\right)\psi_1(\rho,\phi)\right]^*$$

The Eigenvalue Problem of L_z in Two Dimensions (cont.)

To obtain a condition on ψ_1 and ψ_2, we integrate the right side by parts, yielding

$$\begin{aligned}\text{RHS} &= i\,\hbar \int_0^\infty d\rho\,\rho \left[\psi_2(\rho,\phi)\,\psi_1^*(\rho,\phi)\Big|_0^{2\,\pi} - \int_0^{2\,\pi} d\phi\,\psi_1^*(\rho,\phi)\left(\frac{\partial}{\partial\phi}\right)\psi_2(\rho,\phi)\right] \\ &= i\,\hbar \int_0^\infty d\rho\,\rho\,\psi_2(\rho,\phi)\,\psi_1^*(\rho,\phi)\Big|_0^{2\,\pi} + \text{LHS}\end{aligned}$$

We require RHS = LHS for any ψ_1, ψ_2 (not just eigenfunctions), including any possible radial dependence, so we must have

$$\psi(\rho,0) = \psi(\rho,2\,\pi)$$

at any ρ for any ψ. If we impose the constrain on the eigenfunctions, we have

$$1 = e^{2\,\pi\,\frac{i}{\hbar}\,\ell_z}$$

which implies

$$\ell_z = m\,\hbar \qquad\qquad m = 0, \pm 1, \pm 2, \ldots$$

The Eigenvalue Problem of L_z in Two Dimensions (cont.)

One could obtain the same result by requiring that the eigenfunctions be single-valued,

$$\psi(\rho, \phi + 2\,\pi) = \psi(\rho, \phi)$$

which is actually a more restrictive constraint than the one we have applied. The problem is that it is not clear that one should impose this constraint because it could be violated up to a constant phase factor with no physically measurable implications. One really only ought to require that the probability density be single-valued,

$$|\psi(\rho, \phi + 2\,\pi)|^2 = |\psi(\rho, \phi)|^2$$

$$\left| R(\rho)\, e^{\frac{i}{\hbar}\,\ell_z\,\phi} e^{\frac{i}{\hbar}\,2\,\pi\,\ell_z} \right|^2 = \left| R(\rho)\, e^{\frac{i}{\hbar}\,\ell_z\,\phi} \right|^2$$

This results in the requirement $\left| e^{\frac{i}{\hbar}\,2\,\pi\,\ell_z} \right| = 1$, which only implies ℓ_z is real. We already know that from Hermiticity, though; we did not need to require the above condition. Or, put another way, Hermiticity alone implies that the probability density is single-valued. Note also that the lack of single-valuedness has no implications for the action of any operators because it is a position-independent unity-modulus factor that arises.

The Eigenvalue Problem of L_z in Two Dimensions (cont.)

It turns out that, if one considers superposition states of the form

$$\psi(\rho, \phi) = A(\rho)\, e^{\frac{i}{\hbar}\ell_z \phi} + B(\rho)\, e^{\frac{i}{\hbar}\ell_z' \phi}$$

and requires that their probability density always be single-valued, then one can obtain the condition $\ell_z - \ell_z' = m\hbar$, $m = 0, \pm 1, \pm 2, \ldots$. This, combined with the additional fact that the eigenvalues must be symmetric about 0 (because if $\psi(\rho, \phi)$ is an eigenfunction of L_z, then we can complex conjugate the eigenvalue-eigenvector equation to obtain that $\psi(\rho, -\phi)$ ought also be an eigenfunction), implies ℓ_z must either be an integer multiple of $\hbar$ or an odd half-integer multiple of $\hbar$, but one cannot show that only the integer multiple solution holds based on single-valuedness alone.

The Eigenvalue Problem of L_z in Two Dimensions (cont.)

We shall take as our normalized azimuthal eigenfunctions

$$\Phi_m(\phi) = \frac{1}{\sqrt{2\,\pi}}\, e^{i\,m\,\phi} \qquad\qquad m = 0, \pm 1, \pm 2, \ldots$$

They obey the orthonormality condition

$$\int_0^{2\,\pi} d\phi\, \Phi_m^*(\phi)\, \Phi_{m'}(\phi) = \delta_{mm'}$$

The full eigenfunctions are of the form

$$\psi(\rho, \phi) = R(\rho)\, \Phi_m(\phi)$$

There is huge degeneracy for each eigenvalue m because the radial wavefunction is completely unspecified.

Rotationally Invariant Problems in Two Dimensions

It is straightforward to see that, if the potential has no ϕ dependence, then $[H, L_z] = 0$. First, we show that the kinetic energy term always commutes with L_z:

$$\begin{aligned}[P_x^2 + P_y^2, L_z] &= P_x\, P_x\, L_z - L_z\, P_x\, P_x + P_y\, P_y\, L_z - L_z\, P_y\, P_y \\ &= P_x\, [P_x, L_z] + [P_x, L_z]\, P_x + P_y\, [P_y, L_z] + [P_y, L_z]\, P_y \\ &= -i\,\hbar\, P_x\, P_y - i\,\hbar\, P_y\, P_x + i\,\hbar\, P_y\, P_x + i\,\hbar\, P_x\, P_y = 0\end{aligned}$$

Second, when L_z is projected onto the position basis and written in polar coordinates, we see that L_z only has derivatives with respect to ϕ. Therefore, $[L_z, V(\rho)] = 0$ for potentials that have no ϕ dependence and hence $[H, L_z] = 0$.

The Eigenvalue Problem of L_z in Two Dimensions (cont.)

It is therefore useful to solve for simultaneous eigenfunctions of H and L_z to break the degeneracy in L_z (and to of course obtain the eigenfunctions of H classified by their L_z eigenvalue). In polar coordinates with a radial potential $V(\rho)$, the eigenvector-eigenvalue equation for the Hamiltonian is

$$\left[-\frac{\hbar^2}{2\,\mu}\left(\frac{\partial^2}{\partial\rho^2}+\frac{1}{\rho}\,\frac{\partial}{\partial\rho}+\frac{1}{\rho^2}\,\frac{\partial^2}{\partial\phi^2}\right)+V(\rho)\right]\psi_E(\rho,\phi)=E\,\psi_E(\rho,\phi)$$

(We use μ instead of m for the mass to avoid confusion with the L_z eigenvalue index m.) The first term is obtained by rewriting the standard cartesian kinetic energy term in two dimensions in polar coordinates. Doing it by chain rule is quite cumbersome, so we omit the proof here; it can be found in any vector calculus textbook. It should be clear that the third term in the kinetic energy is proportional to L_z^2.

(As an aside, one might ask whether it would be easier to define position operators in polar coordinates, R and Φ, and project directly onto their eigenstates, which we could call $|\rho,\phi\rangle$. The problem is that it is difficult to define a Φ operator in a reasonable way because the ϕ coordinate is not single-valued — multiple values of ϕ corresponds to the same basis element. This problem is discussed in Liboff Problems 9.15 and 9.16 and references therein.)

Guided by $[H,L_z]=0$, let's assume the solution is of the form of an eigenfunction of L_z with eigenvalue ℓ_z, $\psi_{E,m}(\rho,\phi)=R_{E,m}(\rho)\,\Phi_m(\phi)$, and with the form of the radial equation and the energy eigenvalue to be specified by the above differential equation. We shall see why we allow a dependence of R on m below.

The Eigenvalue Problem of L_z in Two Dimensions (cont.)

Inserting the form $\psi_E(\rho,\phi) = R_{E,m}(\rho)\,\Phi_m(\phi)$ into the Hamiltonian's eigenvector-eigenvalue equation yields

$$\left[-\frac{\hbar^2}{2\,\mu}\left(\frac{\partial^2}{\partial\rho^2}+\frac{1}{\rho}\,\frac{\partial}{\partial\rho}-\frac{m^2}{\rho^2}\right)+V(\rho)\right]R_{E,m}(\rho)\,\Phi_m(\phi) = E\,R_{E,m}(\rho)\,\Phi_m(\phi)$$

$\Phi_m(\phi)$ never vanishes, and no derivatives act on it now, so we may cancel it out, and also convert all the radial partial derivatives to total derivatives, leaving the radial equation

$$\left[-\frac{\hbar^2}{2\,\mu}\left(\frac{d^2}{d\rho^2}+\frac{1}{\rho}\,\frac{d}{d\rho}-\frac{m^2}{\rho^2}\right)+V(\rho)\right]R_{E,m}(\rho) = E\,R_{E,m}(\rho)$$

which now depends on and determines only the radial part of the eigenfunction and the eigenvalue E. In general, the eigenvalue E and the radial wavefunction will depend on m because of its presence in the equation. The solution can be determined when one knows the particular form for $V(\rho)$.

Lecture 44:
Rotations and Orbital Angular Momentum in Three Dimensions
The Eigenvalue Problem of L_z and L^2 in Three Dimensions: Differential Equations Method

Date Revised: 2009/02/11

Date Given: 2009/02/09

Classical Rotations in Three Dimensions

One can show (though we will not prove it here, see the Ph106 Lecture Notes), that any rotation in three dimensions can always be decomposed into a "two-dimensional" rotation about a single axis. Hence, we need only consider the extension of our formalism for two-dimensional rotations to allow the axis to point in an arbitrary direction. Let us first just consider rotations about the x or y axes. By analogy to our infinitesimal rotation about the z axis, we may write the form for finite and infinitesimal rotations about the x or y axes:

$$\mathbf{R}_{\theta\hat{x}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\theta & -s_\theta \\ 0 & s_\theta & c_\theta \end{bmatrix} \qquad \mathbf{R}_{\theta\hat{y}} = \begin{bmatrix} c_\theta & 0 & s_\theta \\ 0 & 1 & 0 \\ -s_\theta & 0 & c_\theta \end{bmatrix}$$

$$\mathbf{R}_{\delta\theta\hat{x}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\delta\theta \\ 0 & \delta\theta & 1 \end{bmatrix} = \mathbf{I} + \delta\theta\,\mathbf{M}_x \qquad \mathbf{M}_x \equiv \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{R}_{\delta\theta\hat{y}} = \begin{bmatrix} 1 & 0 & \delta\theta \\ 0 & 1 & 0 \\ -\delta\theta & 0 & 1 \end{bmatrix} = \mathbf{I} + \delta\theta\,\mathbf{M}_y \qquad \mathbf{M}_y \equiv \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

Rotations and Angular Momentum in Three Dimensions (cont.)

Infinitesimal rotations about different axes commute because they are infinitesimal. For example,

$$\begin{aligned}
\mathbf{R}_{\delta\theta_x\hat{x}}\mathbf{R}_{\delta\theta_y\hat{y}} &= (\mathbf{I}+\delta\theta_x\mathbf{M}_x)(\mathbf{I}+\delta\theta_y\mathbf{M}_y) = \mathbf{I}+\delta\theta_x\mathbf{M}_x+\delta\theta_y\mathbf{M}_y+O(\delta\theta)^2 \\
\mathbf{R}_{\delta\theta_y\hat{y}}\mathbf{R}_{\delta\theta_x\hat{x}} &= (\mathbf{I}+\delta\theta_y\mathbf{M}_y)(\mathbf{I}+\delta\theta_x\mathbf{M}_x) = \mathbf{I}+\delta\theta_x\mathbf{M}_x+\delta\theta_y\mathbf{M}_y+O(\delta\theta)^2 \\
&\approx \mathbf{R}_{\delta\theta_x\hat{x}}\mathbf{R}_{\delta\theta_y\hat{y}}
\end{aligned}$$

The generic form for an infinitesimal rotation is therefore

$$\begin{aligned}
\mathbf{R}_{\delta\vec{\theta}} &= \mathbf{I}+\delta\theta_x\mathbf{M}_x+\delta\theta_y\mathbf{M}_y+\delta\theta_z\mathbf{M}_z \equiv \mathbf{I}+\delta\vec{\theta}\cdot\vec{\mathbf{M}} \\
\text{with}\quad \delta\vec{\theta} &= \hat{x}\,\delta\theta_x+\hat{y}\,\delta\theta_y+\hat{z}\,\delta\theta_z \\
\text{and}\quad \vec{\mathbf{M}} &= \hat{x}\,\mathbf{M}_x+\hat{y}\,\mathbf{M}_y+\hat{z}\,\mathbf{M}_z
\end{aligned}$$

The definition of $\vec{\mathbf{M}}$ is purely for the sake of notational convenience. It turns out that $\vec{\mathbf{M}}$ is not a vector, but is actually a third-rank tensor. We will not use this property here, but we refer those who are interested to the Ph106 Lecture Notes.

Rotations and Angular Momentum in Three Dimensions (cont.)

It follows from the above that finite rotations may be written in the form

$$\mathbf{R}_{\vec{\theta}} = \exp(\vec{\theta} \cdot \vec{\mathbf{M}}) = \exp(\theta_x \mathbf{M}_x + \theta_y \mathbf{M}_y + \theta_z \mathbf{M}_z)$$

The fact that any rotation can be written as a two-dimensional rotation about a particular axis is manifest in the above expression. The noncommutativity of finite rotations about different axes is also preserved: even though the ordering of infinitesimal rotations about the different axes does not matter, one finds that it does matter when the power series expansion of the exponential is considered. You can test this very easily by considering $\pi/2$ rotations about $\widehat{x}$ and $\widehat{y}$.

We make a few more useful points about the $\vec{\mathbf{M}}$ matrices. First, an easy-to-remember form for them is

$$(\mathbf{M}_a)_{bc} = -\epsilon_{abc}$$

where $a = 1, 2, 3$ corresponds to $a = x, y, z$ and where ϵ_{abc} is the Levi-Civita symbol of rank 3; it is completely antisymmetric in its indices, which may take on the values $1, 2, 3$. The symbol takes on the value 1 for cyclic permutations of its indices, -1 for anticyclic permutations, and 0 otherwise. It is a third-rank tensor.

Second, the squares of all three of the matrices are similar to the identity matrix:

$$\mathbf{M}_x^2 = -\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{M}_y^2 = -\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{M}_z^2 = -\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Therefore,

$$|\vec{\mathbf{M}}|^2 = \vec{\mathbf{M}} \cdot \vec{\mathbf{M}} = \mathbf{M}_x^2 + \mathbf{M}_y^2 + \mathbf{M}_z^2 = -2\,\mathbf{I}$$

which is a bit strange for the norm of a vector. That happens because $\vec{\mathbf{M}}$ is not really a vector, but is a third-rank tensor.

The matrices satisfy the cyclic commutation relation

$$[\mathbf{M}_a, \mathbf{M}_b] = \sum_c \epsilon_{abc}\mathbf{M}_c \equiv \epsilon_{abc}\mathbf{M}_c \qquad \Longleftrightarrow \qquad \vec{\mathbf{M}} \times \vec{\mathbf{M}} = \vec{\mathbf{M}}$$

Here we have our first encounter with the **Einstein summation convention**, wherein any repeated indices are assumed to be summed over as indicated above. The above relation is a strange identity, indeed, again because $\vec{\mathbf{M}}$ is not a vector but is a third-rank tensor. Since $\vec{\mathbf{M}} \cdot \vec{\mathbf{M}} = -2\,\mathbf{I}$, we have

$$[\mathbf{M}_a, \vec{\mathbf{M}} \cdot \vec{\mathbf{M}}] = 0$$

Finally, just as $\vec{r}^{\,T}\mathbf{M}_z\vec{p} = -\ell_z$, we have in general

$$\vec{\ell} = -\vec{r}^{\,T}\vec{\mathbf{M}}\,\vec{p}$$

Quantum Mechanical Rotations in Three Dimensions

The obvious extension of our quantum two-dimensional and classical three-dimensional rotation formalisms is to recognize that the three quantum angular momentum component operators will generate rotation transformations about their respective axes and that they may be treated as a vector whose inner product with a particular rotation vector will generate that rotation. That is, we have

$$L_x = Y\,P_z - Z\,P_y \qquad L_y = Z\,P_x - X\,P_z \qquad L_z = X\,P_y - Y\,P_x$$
$$\vec{L} = \widehat{x}\,L_x + \widehat{y}\,L_y + \widehat{z}\,L_z$$

The components satisfy the commutation relation

$$[L_a, L_b] = \epsilon_{abc}\,i\,\hbar\,L_c \qquad \Longleftrightarrow \qquad \vec{L}\times\vec{L} = i\,\hbar\,\vec{L}$$

(note, Einstein summation convention used!) just as the $\mathbf{M}_a$ matrices do, up to a factor of $i\,\hbar$. By analogy to our three-dimensional classical and two-dimensional quantum rotation operators, finite rotations are then obtained via

$$T(\vec{\theta}) = \exp\left(-\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{L}\right)$$

Additionally, one can show

$$[R_a, L_b] = \epsilon_{abc}\, i\,\hbar\, R_c \qquad [P_a, L_b] = \epsilon_{abc}\, i\,\hbar\, P_c$$

where R_a are the position component operators X, Y, and Z. (Einstein summation convention, again!)

Just as we calculated $\vec{\mathbf{M}} \cdot \vec{\mathbf{M}}$, we may also calculate L^2,

$$L^2 = L_x^2 + L_y^2 + L_z^2$$

One may verify that

$$[L_a, L^2] = 0$$

which recalls a similar property of $\vec{\mathbf{M}} \cdot \vec{\mathbf{M}}$. So, while no two of the $\{L_a\}$ are simultaneously diagonalizable, one may simultaneously diagonalize any one of the $\{L_a\}$ and the L^2 operator.

The Eigenvector-Eigenvalue Problem of L_z and L^2

Methodology

There are two ways to find the eigenvalues and eigenfunctions of L^2 and L_z:

- **Standard Differential Equations Method**
  Here, we extend the technique we used for two dimensions, obtaining and solving differential equations in ϕ and θ for the eigenfunctions of L^2 and L_z, and finding the allowed values for the eigenvalues by Hermiticity again.
- **Operator Methods**
  One can begin by working in terms of a basis of L^2 and L_z eigenstates and introduce raising and lowering operators along the lines of what was done for the SHO. This lets one study the structure of the eigenvalues of L^2 and L_z without the distraction of the differential equations to determine their position-space representation. The existing of raising and lowering operators also provides a relatively simple means to construct the position-basis representations, again along the lines of what was done for the SHO.

We will pursue both methods. You are probably not yet expert enough in the differential equations method to justify ignoring it completely, though we will not go through all the gore of deriving the Legendre polynomials explicitly. Then we will rely on the operator methodology to better understand the eigenvalue structure and to obtain the full position-basis representations more conveniently.

Differential Equations Method

To rewrite the eigenvalue-eigenvector problems of L_z and L^2 as differential equations, we need to write the action of the operators in a position-basis representation, just as we did for L_z alone in two dimensions. We know the forms for L_x, L_y, and L_z in cartesian coordinates. Putting these three operators on equal footing suggests that the right coordinate system to work in will be spherical coordinates, defined by

$$r = \sqrt{x^2 + y^2 + z^2} \qquad \theta = \arccos \frac{z}{r} \qquad \phi = \arctan \frac{y}{x}$$

We may relabel our cartesian coordinate system position basis using these relations:

$$|x, y, z\rangle = \left| r = \sqrt{x^2 + y^2 + z^2}, \theta = \arccos \frac{z}{r}, \phi = \arctan \frac{y}{x} \right\rangle \tag{14.1}$$

Note: this is not a coordinate transformation, it is only a relabeling of the states that we already know exist.

The Eigenvector-Eigenvalue Problem of L_z and L^2 (cont.)

Let's first project the action of the angular momentum operators onto the position basis:

$$\begin{aligned}\langle x, y, z\,|L_a|\psi\,\rangle &= \langle x, y, z\,|\epsilon_{abc} R_b P_c|\psi\,\rangle \\ &= -i\,\hbar\,\epsilon_{abc} r_b \frac{\partial}{\partial r_c}\psi_q(x, y, z) \\ \langle r, \theta\,\phi\,|L_a|\psi\,\rangle &= -i\,\hbar\,\epsilon_{abc} r_b \frac{\partial}{\partial r_c}\psi_q(r, \theta, \phi)\end{aligned}$$

(Einstein summation convention used) We have skipped the usual steps of inserting completeness to go from the first line to the second line (review Section 5.2 of these notes if you do not recall how to do this.) To go from the second line to the third, we have made use of the equality between cartesian and spherical coordinate system position-basis elements in Equation 14.1 in order to modify the left side. The change to the ride side is a change of the independent variables on which ψ_q depends — it is mathematics, not physics.

The Eigenvector-Eigenvalue Problem of L_z and L^2 (cont.)

Next, we do a change of variables from cartesian to spherical coordinates; again, this is mathematics (calculus), there is no physics involved. It is more tedious than for two dimensions, but you may do it yourself or look it up in a vector calculus text. The result is

$$\langle r, \theta, \phi \, | L_x | \psi \rangle = i \hbar \left(s_\phi \frac{\partial}{\partial \theta} + o_\theta \, c_\phi \frac{\partial}{\partial \phi} \right) \psi_q(r, \theta, \phi)$$

$$\langle r, \theta, \phi \, | L_y | \psi \rangle = i \hbar \left(-c_\phi \frac{\partial}{\partial \theta} + o_\theta \, s_\phi \frac{\partial}{\partial \phi} \right) \psi_q(r, \theta, \phi)$$

$$\langle r, \theta, \phi \, | L_z | \psi \rangle = -i \hbar \frac{\partial}{\partial \phi} \psi_q(r, \theta, \phi)$$

where we introduce $o_\theta = \cot \theta$ and of course $s_\phi = \sin \phi$ and $c_\phi = \cos \phi$.

We will also need the L^2 operator, which is straightforward (though tedious) to calculate from the above:

$$\langle r, \theta, \phi \, | L^2 | \psi \rangle = -\hbar^2 \left(\frac{1}{s_\theta} \frac{\partial}{\partial \theta} s_\theta \frac{\partial}{\partial \theta} + \frac{1}{s_\theta^2} \frac{\partial^2}{\partial \phi^2} \right) \psi_q(r, \theta, \phi)$$

Note that the first $\frac{\partial}{\partial \theta}$ acts on everything to its right including the s_θ factor.

The Eigenvector-Eigenvalue Problem of L_z and L^2 (cont.)

Let us now restrict ourselves to L_z and L^2 alone. Clearly, the process of solving for the L_z eigenfunctions is as before, though now we must allow a dependence on both r and θ. Moreover, the L^2 eigenvalue-eigenvector equation has no dependence on r, so the dependence on r and θ may be separated. So, we may immediately assume

$$\psi_{\alpha,m}(r,\theta,\phi) = R(r)\Theta_{\alpha,m}(\theta)\,\Phi_m(\phi)$$

$$\Phi_m(\phi) = \frac{e^{i\,m\,\phi}}{\sqrt{2\,\pi}} \qquad \ell_z = m\,\hbar \qquad m = 0, \pm 1, \pm 2, \ldots$$

The radial function $R(r)$ is again arbitrary because neither L^2 nor L_z include any r dependence. The polar angle function $\Theta(\theta)$ will depend not just on the L^2 eigenvalue (which we denote for now as α) but also the m eigenvalue because of the ϕ derivative in L^2.

The above form automatically satisfies the L_z eigenvalue-eigenvector equation, which is (projected into the position basis)

$$\langle r,\theta,\phi\,|L_z|\psi_{\alpha,m}\,\rangle = m\,\hbar\,\langle r,\theta,\phi\,|\psi_{\alpha,m}\,\rangle$$

The Eigenvector-Eigenvalue Problem of L_z and L^2 (cont.)

Next, we insert the above form into the L^2 eigenvector-eigenvalue equation projected onto the position basis:

$$\langle r, \theta, \phi \, | L^2 | \psi_{\alpha,m} \rangle = \alpha \langle r, \theta, \phi \, | \psi_{\alpha,m} \rangle$$

We have already calculated the left side, which yields differential operators acting on the position-space wavefunction $\psi_{\alpha,m}(r, \theta, \phi)$. After applying the ϕ derivatives and canceling out the nowhere-vanishing Φ_m, canceling out a radial function $R(r)$ (which may vanish at specific r, but certainly not at all r), moving the eigenvalue to the left side, and replacing the partial derivatives with respect to θ with total derivatives, we obtain

$$-\hbar^2 \left(\frac{1}{s_\theta} \frac{d}{d\theta} s_\theta \frac{d}{d\theta} + \frac{\alpha}{\hbar^2} - \frac{m^2}{s_\theta^2} \right) \Theta_{\alpha,m}(\theta) = 0$$

This is now just a differential equation in θ. Let us change variables to $u = c_\theta$ and define $P_\alpha^m(u = c_\theta) = \Theta_{\alpha,m}$. This yields

$$\left\{ \frac{d}{du} \left[(1 - u^2) \frac{d}{du} \right] + \left[\frac{\alpha}{\hbar^2} - \frac{m^2}{1 - u^2} \right] \right\} P_\alpha^m(u) = 0 \qquad -1 \leq u \leq 1$$

The Eigenvector-Eigenvalue Problem of L_z and L^2 (cont.)

If we set $m = 0$, we obtain

$$\left\{ \frac{d}{du}\left[(1-u^2)\frac{d}{du}\right] + \frac{\alpha}{\hbar^2} \right\} P_\alpha^0(u) = 0$$

which is now an ordinary second-order differential equation with polynomial nonlinear coefficients. You know that the standard solution technique is a series solution.

We will not subject you to the gore of doing the series solution. It suffices to say that one obtains a recursion condition that relates C_{n+2} to C_n in an expansion in powers of u, and that termination results in the requirement

$$\frac{\alpha}{\hbar^2} = \ell(\ell+1) \qquad \ell = 0, 1, 2, \ldots$$

The functions are polynomials containing either even or odd powers of u. They are termed the Legendre polynomials and are denoted by $P_\ell(u)$. There is a formula, called Rodrigues' Formula, that can be used to generate them:

$$P_\ell(u) = \frac{1}{2^\ell\, \ell!} \frac{d^\ell}{du^\ell} \left(u^2 - 1\right)^\ell \tag{14.2}$$

The Eigenvector-Eigenvalue Problem of L_z and L^2 (cont.)

The full equation, with $m \neq 0$, is solved by the associated Legendre polynomials, which can be generated from the formula

$$P_\ell^m(u) = (-1)^m (1-u^2)^{m/2} \frac{d^m}{du^m} P_\ell(u) \qquad 0 \leq m \leq \ell \tag{14.3}$$

where $m \leq \ell$ is enforced by the fact that $P_\ell(u)$ is a ℓth-order polynomial; hence, any derivatives of order $m+1$ or larger simply vanish. We may now write $P_\ell(u)$ as $P_\ell^0(u)$ based on this formula. For $m < 0$, we see that the θ differential equation is unchanged by the sign of m, so we define

$$P_\ell^{-m}(u) = P_\ell^m(u) \qquad 0 \leq m \leq \ell$$

The Eigenvector-Eigenvalue Problem of L_z and L^2 (cont.)

Combining all of this, we thus write our joint eigenfunctions of L^2 and L_z as

$$Y_\ell^m(\theta,\phi) = \sqrt{\frac{2\,\ell+1}{4\,\pi}\frac{(\ell-m)!}{(\ell+m)!}}\,P_\ell^m(u=c_\theta)\,e^{i\,m\,\phi} \tag{14.4}$$

where the prefactor ensures correct normalization when integrated over all solid angles. These functions are known as the spherical harmonics because they are harmonic (sinusoidal) in the spherical coordinate system variables θ and ϕ. The orthonormalization condition (arising from the separate orthonormality of the polar and azimuthal functions) is

$$\int_{-1}^{1} dc_\theta \int_0^{2\,\pi} d\phi\,[Y_\ell^m(\theta,\phi)]^*\,Y_{\ell'}^{m\,\prime}(\theta,\phi) = \delta_{\ell\ell'}\delta_{mm'}$$

The full wavefunction may have any radial dependence as long as its angular dependence is in the form of a spherical harmonic. We summarize our derivation by stating that

$$\langle r,\theta,\phi\,|\psi\rangle = R(r)Y_\ell^m(\theta,\phi) \quad\Longleftrightarrow\quad \begin{array}{l}\langle r,\theta,\phi\,|L^2|\psi\rangle = \ell(\ell+1)\hbar^2\langle r,\theta,\phi\,|\psi\rangle \\ \langle r,\theta,\phi\,|L_z|\psi\rangle = m\,\hbar\langle r,\theta,\phi\,|\psi\rangle\end{array},$$

Lecture 45:
The Eigenvalue Problem of L_z and L^2 in Three Dimensions, ct'd: Operator Method

Date Revised: 2009/02/17

Date Given: 2009/02/11

Operator Method

We've found the eigenfunctions and eigenvalues in the standard pedestrian way. Let's now use some clever operator methods that recall how we used raising and lowering operators to determine the eigenvalues of the SHO without having to explicitly find the eigenfunctions. We shall see that this method leads to a simpler way to find the eigenfunctions too, just as we were able to obtain all the eigenfunctions of the SHO by applying the raising operator in the position basis to the simple Gaussian ground-state wavefunction.

Let's assume we know nothing about the eigenvalue spectrum of L^2 and L_z except that the operators commute so they have simultaneous eigenvectors. Denote an eigenstate of L^2 and L_z with eigenvalues α and β by $|\alpha, \beta\rangle$. That is

$$L^2|\alpha, \beta\rangle = \alpha|\alpha, \beta\rangle \qquad L_z|\alpha, \beta\rangle = \beta|\alpha, \beta\rangle$$

The Eigenvector-Eigenvalue Problem of L_z and L^2 (cont.)

We define angular momentum raising and lowering operators:

$$L_\pm = L_x \pm i\, L_y$$

They are named this way because they satisfy

$$[L_z, L_\pm] = \pm\hbar\, L_\pm$$

so that

$$L_z\,(L_\pm|\alpha,\beta\,\rangle) = (\pm\hbar\, L_\pm + L_\pm\, L_z)\,|\alpha,\beta\,\rangle = (\pm\hbar + \beta)\,(L_\pm|\alpha,\beta\,\rangle)$$

That is, when $|\alpha,\beta\,\rangle$ has L_z eigenvalue β, the state obtained by applying a raising or lowering operator in the state, $L_\pm|\alpha,\beta\,\rangle$, is an eigenvector of L_z with eigenvalue $\beta \pm \hbar$.

The raising and lowering operators commute with L^2,

$$[L^2, L_\pm] = 0$$

so we are assured that $|\alpha,\beta\,\rangle$ and $L_\pm|\alpha,\beta\,\rangle$ have the same eigenvalue α of L^2.

The Eigenvector-Eigenvalue Problem of L_z and L^2 (cont.)

So, our space will break down into subspaces that are eigenspaces of L^2, which will be further decomposed into subspaces that are eigenspaces of L_z. $L_\pm$ moves between these subspaces of a particular L^2 eigenspace. Explicitly, we have

$$L_\pm|\alpha,\beta\rangle = C_\pm(\alpha,\beta)|\alpha,\beta\pm\hbar\rangle$$

We run into the same problem we had with the SHO raising and lowering operators, which is that we so far have no condition that puts a lower or upper limit on the L_z eigenvalue β. Heuristically, it would be unphysical to have $\beta^2 > \alpha$. This can be seen rigorously as follows:

$$\langle\alpha,\beta|\left(L^2 - L_z^2\right)|\alpha,\beta\rangle = \langle\alpha,\beta|\left(L_x^2 + L_y^2\right)|\alpha,\beta\rangle$$

The latter expression is nonnegative because the eigenvalues of L_x^2 and L_y^2 are all nonnegative because the eigenvalues of L_x and L_y are real because they are Hermitian. So we see $\alpha - \beta^2 \geq 0$, or $\alpha \geq \beta^2$ as desired.

The Eigenvector-Eigenvalue Problem of L_z and L^2 (cont.)

So, we require there to be states $|\alpha, \beta_{max}\rangle$ and $|\alpha, \beta_{min}\rangle$ that satisfy

$$L_+|\alpha, \beta_{max}\rangle = 0 \qquad L_-|\alpha, \beta_{min}\rangle = 0$$

where by 0 we mean the null vector, usually referred to as $|0\rangle$, which may be confusing in this situation. We need to rewrite these expressions in terms of L^2 and L_z to further reduce them; let's apply L_- and L_+ to do this:

$$\begin{aligned} L_-L_+|\alpha, \beta_{max}\rangle &= 0 \\ \left(L^2 - L_z^2 - \hbar L_z\right)|\alpha, \beta_{max}\rangle &= 0 \\ \left(\alpha - \beta_{max}^2 - \hbar\beta_{max}\right)|\alpha, \beta_{max}\rangle &= 0 \\ \beta_{max}\left(\beta_{max} + \hbar\right) &= \alpha \end{aligned} \qquad \begin{aligned} L_+L_-|\alpha, \beta_{min}\rangle &= 0 \\ \left(L^2 - L_z^2 + \hbar L_z\right)|\alpha, \beta_{min}\rangle &= 0 \\ \left(\alpha - \beta_{min}^2 + \hbar\beta_{min}\right)|\alpha, \beta_{min}\rangle &= 0 \\ \beta_{min}\left(\beta_{min} - \hbar\right) &= \alpha \end{aligned}$$

which implies

$$\beta_{min} = -\beta_{max}$$

The Eigenvector-Eigenvalue Problem of L_z and L^2 (cont.)

In order for the raising chain begun at β_{min} and the lowering chain begun at β_{max} to terminate, it is necessary that there be a k_+ and k_- such that

$$(L_+)^{k_+ +1} |\alpha, \beta_{min}\rangle \propto |\alpha, \beta_{max}\rangle \qquad (L_-)^{k_- +1} |\alpha, \beta_{max}\rangle \propto |\alpha, \beta_{min}\rangle$$

Therefore

$$\beta_{min} + \hbar k_+ = \beta_{max} \qquad \beta_{max} - \hbar k_- = \beta_{min}$$

So we have

$$k_+ = k_- \equiv k \qquad \beta_{max} - \beta_{min} = \hbar k$$

Since $\beta_{min} = -\beta_{max}$, we then have

$$\beta_{max} = k\,\frac{\hbar}{2} \qquad \alpha = \beta_{max}\,(\beta_{max} + \hbar) = \hbar^2\,\frac{k}{2}\left(\frac{k}{2} + 1\right) \qquad k = 0, 1, 2, \ldots$$

For k even, we recover the allowed eigenvalues we obtained via the differential equation method. The k odd eigenvalues are a different beast, though, and are associated with spin, a degree of freedom that behaves like angular momentum in many ways but is not associated with orbital motion of a particle.

The Eigenvector-Eigenvalue Problem of L_z and L^2 (cont.)

The last point is a very important one: the k odd values arose only from the assumption of the angular momentum operator commutation relations. They did not come from the differential equations, which is what ties all of this to the behavior of spatial wavefunctions; the differential equations method does not permit k odd. This is the source of our statement that the k odd values are not associated with orbital angular momentum. In detail, the restriction to k even comes from the requirement that the wavefunction be single-valued in ϕ, which is required by Hermiticity of L_z. Such a requirement would not hold for a particle spin's z-component operator because there will be no spatial wavefunction to consider. Thus, the above proof tells us which values of k are allowed, and then other restrictions can further reduce the set.

Unlike Shankar, who gives a bit more detailed of a hint at what is meant by spin, we will delay discussion until we have time to do it thoroughly. For now it is not important to have a physical picture of the states that result in half-integral values of L_z.

The Eigenvector-Eigenvalue Problem of L_z and L^2 (cont.)

Given that the spectrum of eigenvalues we have derived is more general than just orbital angular momentum L, we will follow standard notation and use J instead of L. We will denote the eigenvalues as follows:

- We will denote by j the value of $k/2$. j may take on any nonnegative integral or half-integral value.
- The J^2 eigenvalue is $\alpha = \hbar^2 j(j+1)$. However, we will replace α in $|\alpha, \beta\rangle$ by j for brevity.
- The J_z eigenvalue β can take on values from $-j\hbar$ to $j\hbar$ in steps of size $\hbar$. We define $m = \beta/\hbar$. We will replace β in $|\alpha, \beta\rangle$ by m for consistency with the notation we developed via the differential equation method.

Therefore, simultaneous eigenstates of J^2 and J_z will be denoted by $|j, m\rangle$ and will have J^2 eigenvalue $\alpha = \hbar^2 j(j+1)$ and J_z eigenvalue $\beta = m\hbar$.

The Eigenvector-Eigenvalue Problem of L_z and L^2 (cont.)

Summary

Let us take a step back and see what we have done and where we should go. What we have done:

- We are considering problems in two or three spatial dimensions in cylindrical and spherical coordinates with an eye toward working with Hamiltonians that are invariant under rotations and hence depend on the cylindrical coordinate ρ or the radial coordinate r.
- Since a continuous symmetry transformation of a Hamiltonian derives from a generator operator that commutes with the Hamiltonian, we knew it would be useful to find the generator and its eigenvalues and eigenvectors to help us reduce or solve the eigenvector-eigenvalue problem of the full Hamiltonian.
- This led us to write explicit forms for $\vec{L}$ and L^2 and to obtain their eigenvectors and eigenfunctions, both in the position basis and in the more natural basis of their eigenstates.
- We have thus been able to organize the Hilbert space into subspaces of specific values of the angular momentum magnitude.

The Eigenvector-Eigenvalue Problem of L_z and L^2 (cont.)

We have two important tasks left:

- To understand the full structure of the Hilbert space in terms of the eigenstates of J^2 and J_z; *i.e.*, let's write down explicit forms for all the operators we have considered: J_x, J_y, J_+, J_- and rotation operators.
- To understand the connection between the $\{|j, m\rangle\}$ basis and the position basis eigenstates — essentially, to show that we can obtain the position basis eigenstates from the structure of the Hilbert space in terms of the $\{|j, m\rangle\}$ basis.

We consider these tasks next.

Lecture 46:
Angular Momentum Operators in the $\{|j, m\rangle\}$ Basis

Date Revised: 2009/02/23
Date Given: 2009/02/13

Angular Momentum Operators in the $|j,m\rangle$ Basis

To evaluate explicitly the non-diagonal angular momentum operators J_x, J_y, J_+, and J_- in the $\{|j,m\rangle\}$ basis, we need to determine the coefficient $C(\alpha,\beta)$ in the relation

$$J_\pm|\alpha,\beta\rangle = C_\pm(\alpha,\beta)|\alpha,\beta\pm\hbar\rangle$$

which we should now rewrite as

$$J_\pm|j,m\rangle = C_\pm(j,m)|j,m+1\rangle$$

Again, we use our SHO example as a guide for how to determine this coefficient; there, we required each eigenstate to be normalized, which puts a constraint on the C's. So:

$$\begin{aligned}
1 = \langle j,m\pm1\,|j,m\pm1\rangle &= |C_\pm(j,m)|^{-2}\,|J_\pm|j,m\rangle|^2 \\
&= |C_\pm(j,m)|^{-2}\,\langle j,m\,|J_\mp J_\pm|j,m\rangle \\
&= |C_\pm(j,m)|^{-2}\,\langle j,m\,|\left(J^2-J_z^2\mp\hbar J_z\right)|j,m\rangle \\
&= |C_\pm(j,m)|^{-2}\left[j(j+1)-m^2\mp m\right]\hbar^2 \\
\Longrightarrow |C_\pm(j,m)|^2 &= \hbar^2(j\mp m)(j\pm m+1)
\end{aligned}$$

Operators in the $\{|j,m\rangle\}$ Basis (cont.)

We discard the phase freedom and take

$$C_{\pm}(j,m) = \hbar\sqrt{(j \mp m)(j \pm m + 1)}$$

So we are left with

$$J_{\pm}|j,m\rangle = \hbar\sqrt{(j \mp m)(j \pm m + 1)}|j,m \pm 1\rangle \tag{14.5}$$

We have the expected result that $J_{\pm}$ annihilates $|j,\pm j\rangle$. Note that $J_{\pm}$ do not change j, only m. That is, the matrix elements of $J_{\pm}$ between two states $|j,m\rangle$ and $|j',m'\rangle$ are only nonzero if $j = j'$: $J_{\pm}$ do not connect states of different j.

Operators in the $\{|j,m\rangle\}$ Basis (cont.)

So, we have obtained all the simultaneous eigenstates of J^2 and J_z, properly normalized, and we also know the action of $J_\pm$, which finally lets us write the matrix elements of J_x and J_y in the $\{|j,m\rangle\}$ basis:

$$\begin{aligned}\langle j',m'|J_x|j,m\rangle &= \langle j',m'|\left(\frac{J_+ + J_-}{2}\right)|j,m\rangle \\ &= \frac{\hbar}{2}\Big[\delta_{j,j'}\delta_{m',m+1}\sqrt{(j-m)(j+m+1)} \\ &\qquad +\delta_{j,j'}\delta_{m',m-1}\sqrt{(j+m)(j-m+1)}\Big] \end{aligned} \tag{14.6}$$

Similarly,

$$\begin{aligned}\langle j',m'|J_y|j,m\rangle &= \langle j',m'|\left(\frac{J_+ - J_-}{2i}\right)|j,m\rangle \\ &= \frac{\hbar}{2i}\Big[\delta_{j,j'}\delta_{m',m+1}\sqrt{(j-m)(j+m+1)} \\ &\qquad -\delta_{j,j'}\delta_{m',m-1}\sqrt{(j+m)(j-m+1)}\Big] \end{aligned} \tag{14.7}$$

As with $J_\pm$, the matrix elements of J_x and J_y between two states $|j,m\rangle$ and $|j',m'\rangle$ are only nonzero if $j=j'$: J_x and J_y do not connect states of different j.

Operators in the $\{|j,m\rangle\}$ Basis (cont.)

Let's consider what the matrix representations of these various operators *look like* in the $|j,m\rangle$ basis. J^2 and J_z are completely diagonal because the basis consists of their eigenvectors. J_x, J_y, and $J_\pm$ will be block diagonal: since these operators do not change j, their matrix elements between states of different j vanish (as noted above: they always have $\delta_{j,j'}$). The block diagonal nature of the matrix representations is clear in the forms written out in Shankar Equations 12.5.22, 12.5.23, and 12.5.24, which we reproduce here, with the ordering of the basis elements being $|0,0\rangle$, $|\frac{1}{2},\frac{1}{2}\rangle$, $|\frac{1}{2},-\frac{1}{2}\rangle$, $|1,1\rangle$, $|1,0\rangle$, $|1,-1\rangle$ (the $^{(0)\oplus(1/2)\oplus(1)}$ notation tells us which subspace these operators are restricted to):

$$J_z^{(0)\oplus(1/2)\oplus(1)} \xleftrightarrow[|j,m\rangle]{} \hbar \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

$$[J^2]^{(0)\oplus(1/2)\oplus(1)} \xleftrightarrow[|j,m\rangle]{} \hbar^2 \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{3}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{3}{4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$$

Operators in the $\{|j,m\rangle\}$ Basis (cont.)

$$J_+^{(0)\oplus(1/2)\oplus(1)} \xleftrightarrow[|j,m\rangle]{} \hbar \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$J_-^{(0)\oplus(1/2)\oplus(1)} \xleftrightarrow[|j,m\rangle]{} \hbar \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & 0 \end{bmatrix}$$

$$J_x^{(0)\oplus(1/2)\oplus(1)} \xleftrightarrow[|j,m\rangle]{} \frac{\hbar}{2} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{2} & 0 & \sqrt{2} \\ 0 & 0 & 0 & 0 & \sqrt{2} & 0 \end{bmatrix}$$

$$J_y^{(0)\oplus(1/2)\oplus(1)} \xleftrightarrow[|j,m\rangle]{} \frac{i\hbar}{2} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{2} & 0 & -\sqrt{2} \\ 0 & 0 & 0 & 0 & \sqrt{2} & 0 \end{bmatrix}$$

Operators in the $\{|j,m\rangle\}$ Basis (cont.)

This block diagonal nature recalls recalls the idea of direct sum spaces (Section 3.4): we may write the entire vector space as a sum over all the possible j subspaces:

$$\mathbb{V} = \mathbb{V}^{(0)} \oplus \mathbb{V}^{(1/2)} \oplus \mathbb{V}^{(1)} \oplus \mathbb{V}^{(3/2)} \oplus \mathbb{V}^{(2)} \oplus \dots$$

These are just the degenerate subspaces of the J^2 operator. If we restrict to orbital angular momentum, we know the half-integer j values are not allowed, leaving us with

$$\mathbb{V} = \mathbb{V}^{(0)} \oplus \mathbb{V}^{(1)} \oplus \mathbb{V}^{(2)} \oplus \dots$$

But this space is a subspace of the more generic one, so there is no harm in discussing the generic version and then specializing to integer j for particular problems.

Operators in the $\{|j,m\rangle\}$ Basis (cont.)

Let us write out the matrix representations of a single subspace too, the $j = 1$ subspace, which we denote by $\mathbb{V}^{(1)}$ and which we will indicate by putting a $^{(1)}$ subscript on the operators. We have

$$[J^2]^{(1)} \underset{|j,m\rangle}{\longleftrightarrow} 2\hbar^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad J_z^{(1)} \underset{|j,m\rangle}{\longleftrightarrow} \hbar \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

$$J_+^{(1)} \underset{|j,m\rangle}{\longleftrightarrow} \hbar\sqrt{2} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad J_-^{(1)} \underset{|j,m\rangle}{\longleftrightarrow} \hbar\sqrt{2} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$J_x^{(1)} \underset{|j,m\rangle}{\longleftrightarrow} \frac{\hbar}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad J_y^{(1)} \underset{|j,m\rangle}{\longleftrightarrow} \frac{i\hbar}{\sqrt{2}} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

One important warning: just because the $\mathbb{V}^{(1)}$ subspace is three-dimensional, do not misinterpret this as implying it works on the $\mathbb{R}^3$ space of Cartesian vectors in three spatial dimensions. First, $\mathbb{V}^{(1)}$ is isomorphic to $\mathbb{C}^3$, not $\mathbb{R}^3$, because one is allowed to have complex coefficients for QM Hilbert space vectors. Second, even if that were not a problem, it does not hold true that, for example, $|1,1\rangle$ is the same as $\widehat{x}$. We will see the connection to three-dimensional Cartesian vectors later.

Operators in the $\{|j,m\rangle\}$ Basis (cont.)

That is, we have:

- $[J^2]^{(1)}$ is diagonal and the $\mathbb{V}^{(1)}$ subspace is degenerate for it.
- $J_z^{(1)}$ is diagonal and it breaks the degeneracy in the $\mathbb{V}^{(1)}$ subspace. We have put the eigenstates in the order $|1,1\rangle$, $|1,0\rangle$, $|1,-1\rangle$ as is indicated by the ordering of the J_z eigenvalues.
- $J_+^{(1)}$ and $J_-^{(1)}$ are not diagonal, and the former only connects a particular m to $m+1$ and the latter only connects a particular m to $m-1$. $J_+^{(1)}$ annihilates $|1,1\rangle$ and can never yield $|1,-1\rangle$ and similarly $J_-^{(1)}$ annihilates $|1,-1\rangle$ and can never yield $|1,1\rangle$.
- $J_x^{(1)}$ and $J_y^{(1)}$ are not diagonal, but we see that they connect states that differ in m by one unit.

Lecture 47:
Rotation Operators in the $\{|j, m\rangle\}$ Basis

Date Revised: 2009/02/18
Date Given: 2009/02/18

Rotation Operators in the $\{|j,m\rangle\}$ Basis

The J_k operator is the generator of rotations about the k axis, and any rotation can be written in terms of the $\{J_k\}$ as $T(\vec{\theta}) = \exp(-(i/\hbar)\,\vec{\theta}\cdot\vec{J})$. (This holds for odd-half-integer j also, though we have not proven it yet.) Any product of two matrices that are block diagonal in the same manner will be block diagonal too, so we see that the rotation operator will be block diagonal in the same way as the J_k. So we see that rotations do not mix states of different j; they cannot change j.

Just as we denoted above the restriction of the $\vec{J}$ operator to the $\mathbb{V}^{(j)}$ subspace by $\vec{J}^{(j)}$, we can denote the restriction of the $T(\vec{\theta})$ operator to the $\mathbb{V}^{(j)}$ subspace by $T^{(j)}(\vec{\theta})$.

Operators in the $\{|j,m\rangle\}$ Basis (cont.)

If we again look at the $\mathbb{V}^{(0)} \oplus \mathbb{V}^{(1/2)} \oplus \mathbb{V}^{(1)}$ subspace, we can see how the $T^{(j)}(\vec{\theta})$ fit into the larger $T(\vec{\theta})$ operator:

$$T^{(0)\oplus(1/2)\oplus(1)}(\vec{\theta}) \xleftrightarrow[|j,m\rangle]{}$$

$$\begin{bmatrix}
\left[T^{(0)}(\vec{\theta})\right]_{11} & 0 & 0 & 0 & 0 & 0 \\
0 & \left[T^{(1/2)}(\vec{\theta})\right]_{11} & \left[T^{(1/2)}(\vec{\theta})\right]_{12} & 0 & 0 & 0 \\
0 & \left[T^{(1/2)}(\vec{\theta})\right]_{21} & \left[T^{(1/2)}(\vec{\theta})\right]_{22} & 0 & 0 & 0 \\
0 & 0 & 0 & \left[T^{(1)}(\vec{\theta})\right]_{11} & \left[T^{(1)}(\vec{\theta})\right]_{12} & \left[T^{(1)}(\vec{\theta})\right]_{13} \\
0 & 0 & 0 & \left[T^{(1)}(\vec{\theta})\right]_{21} & \left[T^{(1)}(\vec{\theta})\right]_{22} & \left[T^{(1)}(\vec{\theta})\right]_{23} \\
0 & 0 & 0 & \left[T^{(1)}(\vec{\theta})\right]_{31} & \left[T^{(1)}(\vec{\theta})\right]_{32} & \left[T^{(1)}(\vec{\theta})\right]_{33}
\end{bmatrix}$$

Operators in the $\{|j,m\rangle\}$ Basis (cont.)

where

$$\left[T^{(0)}(\vec{\theta})\right]_{11} = \langle 0,0\,|\,T^{(0)}(\vec{\theta})|0,0\,\rangle$$

$$\left[T^{(1/2)}(\vec{\theta})\right]_{11} = \left\langle \frac{1}{2},\frac{1}{2}\,\middle|\,T^{(1/2)}(\vec{\theta})\,\middle|\,\frac{1}{2},\frac{1}{2}\right\rangle \qquad \left[T^{(1/2)}(\vec{\theta})\right]_{12} = \left\langle \frac{1}{2},\frac{1}{2}\,\middle|\,T^{(1/2)}(\vec{\theta})\,\middle|\,\frac{1}{2},-\frac{1}{2}\right\rangle$$

$$\left[T^{(1/2)}(\vec{\theta})\right]_{21} = \left\langle \frac{1}{2},-\frac{1}{2}\,\middle|\,T^{(1/2)}(\vec{\theta})\,\middle|\,\frac{1}{2},\frac{1}{2}\right\rangle \qquad \left[T^{(1/2)}(\vec{\theta})\right]_{22} = \left\langle \frac{1}{2},-\frac{1}{2}\,\middle|\,T^{(1/2)}(\vec{\theta})\,\middle|\,\frac{1}{2},\frac{1}{2}\right\rangle$$

$$\left[T^{(1)}(\vec{\theta})\right]_{11} = \langle 1,1\,|\,T^{(1)}(\vec{\theta})|1,1\,\rangle \qquad \left[T^{(1)}(\vec{\theta})\right]_{12} = \langle 1,1\,|\,T^{(1)}(\vec{\theta})|1,0\,\rangle$$

$$\left[T^{(1)}(\vec{\theta})\right]_{13} = \langle 1,1\,|\,T^{(1)}(\vec{\theta})|1,-1\,\rangle$$

$$\left[T^{(1)}(\vec{\theta})\right]_{21} = \langle 1,0\,|\,T^{(1)}(\vec{\theta})|1,1\,\rangle \qquad \left[T^{(1)}(\vec{\theta})\right]_{22} = \langle 1,0\,|\,T^{(1)}(\vec{\theta})|1,0\,\rangle$$

$$\left[T^{(1)}(\vec{\theta})\right]_{23} = \langle 1,0\,|\,T^{(1)}(\vec{\theta})|1,-1\,\rangle$$

$$\left[T^{(1)}(\vec{\theta})\right]_{31} = \langle 1,-1\,|\,T^{(1)}(\vec{\theta})|1,1\,\rangle \qquad \left[T^{(1)}(\vec{\theta})\right]_{32} = \langle 1,-1\,|\,T^{(1)}(\vec{\theta})|1,0\,\rangle$$

$$\left[T^{(1)}(\vec{\theta})\right]_{33} = \langle 1,-1\,|\,T^{(1)}(\vec{\theta})|1,-1\,\rangle$$

Operators in the $\{|j,m\rangle\}$ Basis (cont.)

We may write explicit forms for the rotation operators in the $\mathbb{V}^{(j)}$ subspaces using the subspace-restricted version of $\vec{J}$. In general, we define

$$T^{(j)}(\vec{\theta}) = \exp\left(-\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{J}^{(j)}\right) = \sum_{n=0}^{\infty}\frac{1}{n!}\left(-\frac{i}{\hbar}\,\theta\right)^n\left(\widehat{\theta}\cdot\vec{J}^{(j)}\right)^n$$

This may seem difficult to evaluate, but it turns out not to be because one can show that $\left(\widehat{\theta}\cdot\vec{J}^{(j)}\right)^n$ for $n > 2j$ can be written as a linear combination of the first $2j$ powers of $\widehat{\theta}\cdot\vec{J}^{(n)}$. The $j = 1$ case is similar to the way in which the generators of classical rotations, the $\vec{\mathbf{M}}$ matrices, satisfied $\mathbf{M}_k^{2n} = (-1)^n\,\mathbf{M}_k^2$ so that only $\mathbf{M}_k$ and $\mathbf{M}_k^2$ were unique and independent.

Operators in the $\{|j,m\rangle\}$ Basis (cont.)

Therefore, we may write

$$T^{(j)}(\vec{\theta}) = \sum_{n=0}^{2j} f_n(\theta) \left(\widehat{\theta}\cdot\vec{J}^{(j)}\right)^n \tag{14.8}$$

Specific examples are:

$$T^{(0)}(\vec{\theta}) = I^{(0)} \tag{14.9}$$

$$T^{(1/2)}(\vec{\theta}) = \cos\frac{\theta}{2} + 2\,\frac{i}{\hbar}\,\widehat{\theta}\cdot\vec{J}^{(1/2)}\,\sin\frac{\theta}{2} \tag{14.10}$$

$$T^{(1)}(\theta_k\widehat{k}) = \left[I^{(1)} + \left(c_{\theta_k} - 1\right)\left(\frac{J_k^{(1)}}{\hbar}\right)^2\right] - i\,s_{\theta_k}\left(\frac{J_k^{(1)}}{\hbar}\right) \tag{14.11}$$

where k runs over x, y, z and $\widehat{k}$ is the unit vector for the corresponding direction. Note that there is no simple form for the $j = 1$ case for arbitrary $\widehat{\theta}$ because of the fact that one must necessarily end up with higher than linear powers of $\vec{J}$, so noncommutativity of the J_i becomes a problem.

Operators in the $\{|j,m\rangle\}$ Basis (cont.)

The fact that the $T(\vec{\theta})$ operators are block diagonal in the same was as the $\vec{J}$ operators leads us to call the $\mathbb{V}^{(j)}$ invariant subspaces: rotations may mix up members of a particular subspace, but they never send a subspace member out of that subspace. "Invariant" is perhaps somewhat misleading because it suggests that rotations have no effect at all on these subspaces. A better term might be "closed" subspaces.

We also term these subspaces irreducible because they contain no smaller invariant subspaces. Shankar offers a detailed proof of this, but, it is rather easy to see the irreducibility from the structure of the matrix representations of $J_x^{(j)}$ and $J_y^{(j)}$. $J_x^{(j)}$ always connects a particular state $|j,m\rangle$ to $|j,m\pm 1\rangle$. A rotation about x will always result in $2j$ nontrivial powers of $J_x^{(j)}$ as we explained above. Since $J_x^{(j)}$ connects $|j,m\rangle$ to $|j,m\pm 1\rangle$, $2j$ powers of $J_x^{(j)}$ will connect $|j,m\rangle$ to all possible $|j,m'\rangle$ since there are at most $2j$ other $|j,m'\rangle$. There may be values of the rotation angle for which these connections vanish, but that will not happen in general. Hence, there is no closed subspace of $\mathbb{V}^{(j)}$ that is smaller than $\mathbb{V}^{(j)}$.

We note that irreducibility of the invariant subspaces is equivalent to saying that they cannot be made "more" block diagonal – *i.e.*, that the blocks cannot be made smaller. We can see this by realizing that, if the block for $\mathbb{V}^{(j)}$ could be made smaller, then the subblocks would indicate the subspaces of $\mathbb{V}^{(j)}$ that are invariant (closed) under rotations, which we have just concluded can be no smaller than $\mathbb{V}^{(j)}$ itself.

Operators in the $\{|j,m\rangle\}$ Basis (cont.)

The block diagonal form of the rotation operators we have obtained is termed an irreducible matrix representation of the rotation operators because it cannot be further block diagonalized; equivalently, because all the invariant subspaces are irreducible. The corresponding block diagonal form of the $\vec{J}$ and J^2 operators is termed an irreducible representation of those operators for the same reason, though the idea of "invariance" does not realy apply because these operators are not performing a unitary transformation. Certainly, though, the idea of "closed" subspaces does apply and suffices here.

Operators in the $\{|j,m\rangle\}$ Basis (cont.)

Relation to the Degenerate Eigensubspaces of Rotationally Invariant Hamiltonians

Let's consider Hamiltonians that satisfy $[H,\vec{J}]=0$. We term these rotationally invariant because this condition implies that a rotation transformation about any axis is a symmetry transformation of H. Note that the condition $[H,\vec{J}]=0$ implies $[H,J^2]=0$. Therefore, our work on eigenvectors and eigenvalues of J^2 and J_z applies.

We can see that the invariant subspaces of the rotation operator, which are also the closed subspaces of the $\vec{J}$ and J^2 operators, must be degenerate eigensubspaces of a rotationally invariant H. That is, all elements of a subspace $\mathbb{V}^{(j)}$ are eigenstates of H with the same eigenvalue E. This is not obvious just from $[H,J^2]=0$, and $[H,J_z]=0$. Those commutation relations imply that eigenstates of H are eigenstates of J^2 and J_z and vice versa. But the commutators imply nothing about whether the eigenvalues of H, J^2, and J_z are related.

We may prove this point about eigensubspaces of H by realizing that our rotational invariance condition includes $[H,J_x]=0$ and $[H,J_y]=0$, which then implies $[H,J_\pm]=0$. If $|j,m\rangle$ is an eigenstate of H with eigenvalue E — which is implied by $[H,J^2]=0$ and $[H,J_z]=0$ — then $[H,J_\pm]=0$ implies that $J_\pm^k|j,m\rangle$ is an eigenstate of H with the same eigenvalue E (unless $J_\pm^k|j,m\rangle=0$). One can reach any of the $\{|j,m\rangle\}$ in $\mathbb{V}^{(j)}$ using enough powers of $J_\pm$, so all the $\{|j,m\rangle\}$ in $\mathbb{V}^{(j)}$ must also be eigenstates of H of energy E. Since the $\{|j,m\rangle\}$ span the $\mathbb{V}^{(j)}$ subspace, $\mathbb{V}^{(j)}$ is thus a degenerate eigensubspace of H with eigenvalue E.

Lecture 48:
Relation between $|j, m\rangle$ Basis and Position Basis Eigenstates

Date Revised: 2009/02/20
Date Given: 2009/02/20

Relation between the $|j, m\rangle$ Basis and Position Basis Eigenstates

We have established the "structure" of the eigenstates and eigenvalues of J^2 and J_z via operator methods. The same structure came out of the "differential equation" method of finding the eigenstates of L^2 and L_z in the position basis, though this structure was far less obvious. Now, let's connect the two by showing how we can use our operator results to derive the position-basis representation of the eigenstates $\{|\ell, m\rangle\}$. This will be very similar to the way we made the analogous connection for the SHO.

We remark that working in the position basis forces us to specialize from J to L because the existence of a position basis with particular matrix elements for the $\vec{L}$ and L^2 operators in that basis is specific to orbital angular momentum. Such a basis simply does not exist for spin, as we shall see from how we define spin.

Relation between $|j, m\rangle$ Basis and Position Basis Eigenstates (cont.)

When we did this for the SHO, we began with the fact that the lowering operator annihilates the ground state. Here, we have that the raising operator annihilates $|\ell, \ell\rangle$ and the lowering operator annihilates $|\ell, -\ell\rangle$. We shall see that we need to us both relations, so let's begin with both:

$$L_{\pm}|\ell, \pm\ell\rangle = 0$$

Let's project onto the position basis:

$$\langle r, \theta, \phi\,|\,(L_x \pm i\,L_y)\,|\ell, \pm\ell\rangle = \langle r, \theta, \phi\,|0\rangle = 0$$

We calculated the above matrix elements for L_x and L_y in Section 13.5 when we began the "differential equations" method, so we use those results to obtain

$$\pm\hbar\, e^{i\phi}\left(\frac{\partial}{\partial\theta} \pm i\, o_\theta \frac{\partial}{\partial\phi}\right)\psi_\ell^{\pm\ell}(r, \theta, \phi) = 0$$

Relation between $|j, m\rangle$ Basis and Position Basis Eigenstates (cont.)

Since we know $\psi_\ell^{\pm\ell}(r, \theta, \phi)$ must be an eigenstate of L_z with eigenvalue $\pm\ell\,\hbar$, we know the solution must be of the form

$$\psi_\ell^{\pm\ell}(r, \theta, \phi) = U_\ell^{\pm\ell}(r, \theta)\, e^{\pm i\,\ell\,\phi}$$

Inserting this, we obtain

$$\left(\frac{d}{d\theta} - \ell\, o_\theta\right) U_\ell^{\pm\ell}(r, \theta) = 0$$

where we canceled out the nonvanishing $e^{\pm i\phi}$ and $e^{\pm i\,\ell\,\phi}$ factors. This is integrable:

$$\frac{dU_\ell^{\pm\ell}}{U_\ell^{\pm\ell}} = \ell\,\frac{d(s_\theta)}{s_\theta}$$

$$U_\ell^{\pm\ell} = R(r)\,(s_\theta)^\ell$$

Relation between $|j, m\rangle$ Basis and Position Basis Eigenstates (cont.)

The angular part of $\psi_\ell^{\pm\ell}$ ought to be the spherical harmonic $Y_\ell^{\pm\ell}$ (Equation 14.4), after we correct for normalization and follow the same sign convention as we used before. It is:

$$Y_\ell^{\pm\ell}(\theta, \phi) = (-1)^\ell \sqrt{\frac{(2\,l+1)!}{4\,\pi}} \frac{1}{2^\ell\,\ell!} (s_\theta)^\ell\, e^{\pm i\,\ell\,\phi}$$

That is, $Y_\ell^{\pm\ell}$ has the same θ and ϕ dependences as the solutions we found to the annihilation conditions.

Relation between $|j, m\rangle$ Basis and Position Basis Eigenstates (cont.)

We of course obtain the position-basis representations by using the lowering and raising operators on the $|\ell, \pm\ell\rangle$ states:

$$
\begin{aligned}
&\psi_\ell^{\pm m}(r, \theta, \phi) \\
&= \langle r, \theta, \phi \,|\ell, \pm m\rangle \\
&= \langle r, \theta, \phi \,|L_\mp^{(\ell - m)}|\ell, \pm\ell\rangle \\
&= \left[\mp\hbar e^{\mp i\phi}\left(\frac{\partial}{\partial\theta} \mp i\, o_\theta \frac{\partial}{\partial\phi}\right)\right]^{(\ell-m)} R(r)\, Y_\ell^{\pm\ell}(\theta, \phi) \\
&= R(r)\,(-1)^\ell(\pm 1)^m \sqrt{\frac{2\,\ell + 1}{4\,\pi}\,\frac{(\ell + m)!}{(\ell - m)!}}\, e^{\pm\, i\, m\, \phi}\,(s_\theta)^{-m}\left(\frac{d}{d(c_\theta)}\right)^{\ell - m}(s_\theta)^{2\,\ell} \\
&= R(r)\, Y_\ell^{\pm m}
\end{aligned}
$$

where we recognize Rodrigues' formula (Equation 14.2) and the formula for the associated Legendre polynomials (Equation 14.3) in the penultimate step (remember, $u = c_\theta$, $s_\theta^2 = u^2 - 1$, $\frac{d}{du} = \frac{d}{d(c_\theta)}$). We recover the spherical harmonics completely.

We note that we used the annihilation conditions on both $|\ell, \pm\ell\rangle$ simply to make explicit the symmetry between m and $-m$ in this procedure; lowering Y_ℓ^0 would have appeared to have broken this symmetry.

Lecture 49:
Rotationally Invariant Problems in Three Dimensions

Date Revised: 2009/02/23
Date Given: 2009/02/23

Rotationally Invariant Problems in Three Dimensions

Rotational invariance in three spatial dimensions means that H is invariant under a rotation about *any* axis. That is, we need $[H, \vec{L}] = 0$. While this condition yields $[H, L^2] = 0$, this latter condition is not enough to meet our definition of rotational invariance.

It is easy to see that $[T, \vec{L}] = 0$ where T is the kinetic energy operator, $T = (P_x^2 + P_y^2 + P_z^2)/2\,m$. We already know $[P_x^2 + P_y^2, L_z] = 0$ from our discussion of rotationally invariant problems in two dimensions. It also easy to see $[P_z, L_z] = 0$ because $L_z = X\,P_y - Y\,P_x$. So $[T, L_z] = 0$. The same kinds of arguments, with the coordinates permuted cyclically, work for $[T, L_x]$ and $[T, L_y]$.

Rotationally Invariant Problems in Three Dimensions (cont.)

Therefore, in order for $[H, \vec{L}] = 0$, we need $[V, \vec{L}] = 0$, which is equivalent to $[V, T(\vec{\theta})] = 0$ for any $\vec{\theta}$. Let's project this equation onto the position basis:

$$
\begin{aligned}
&[V, T(\vec{\theta})] = 0 \\
&V\, T(\vec{\theta}) - T(\vec{\theta})\, V = 0 \\
&T^{\dagger}(\vec{\theta})\, V\, T(\vec{\theta}) = V \\
&\langle x = u_1, y = v_1, z = w_1 \,|\, T^{\dagger}(\vec{\theta})\, V\, T(\vec{\theta}) | x = u_2, y = v_2, z = w_2 \,\rangle \\
&\qquad\qquad = \langle x = u_1, y = v_1, z = w_1 \,|\, V | x = u_2, y = v_2, z = w_2 \,\rangle \\
&\langle x' = u_1, y' = v_1, z' = w_1 \,|\, V | x' = u_2, y' = v_2, z' = w_2 \,\rangle \\
&\qquad\qquad = \langle x = u_1, y = v_1, z = w_1 \,|\, V | x = u_2, y = v_2, z = w_2 \,\rangle \\
&V(x' = u_1, y' = v_1, z' = w_1)\, \delta(u_1 - u_2)\, \delta(v_1 - v_2)\, \delta(z_1 - z_2) \\
&\qquad\qquad = V(x = u_1, y = v_1, z = w_1)\, \delta(u_1 - u_2)\, \delta(v_1 - v_2)\, \delta(z_1 - z_2) \\
&V(x' = u_1, y' = v_1, z' = w_1) = V(x = u_1, y = v_1, z = w_1)
\end{aligned}
$$

The only way for the potential's functional dependence on the primed and unprimed coordinates to be the same for any choice of $\vec{\theta}$ is for the potential to be a function of r only: you know this from classical mechanics. So we restrict to potentials that depend on radius alone, $V = V(r)$.

Rotationally Invariant Problems in Three Dimensions (cont.)

The eigenvector-eigenvalue equation for the Hamiltonian with such a potential in three dimensions in spherical coordinates is

$$\left\{-\frac{\hbar^2}{2\,\mu}\left[\frac{1}{r^2}\,\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{1}{r^2 s_\theta}\,\frac{\partial}{\partial\theta}\left(s_\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{r^2 s_\theta^2}\,\frac{\partial^2}{\partial\phi^2}\right]+V(r)\right\}\psi_E(r,\theta,\phi)$$
$$= E\,\psi_E(r,\theta,\phi)$$

Referring back to Section 13.5 where we wrote down the differential equation for $L^2\psi$, we see that the angular terms here are $-L^2/r^2$. The equation will thus simplify greatly when we assume a solution that is an eigenvector of L^2. We should also require it to be an eigenvector of L_z since the Hamiltonian commutes with L_z. So we assume

$$\psi_E(r,\theta,\phi) = R_{E,\ell}(r)\,Y_\ell^m(\theta,\phi)$$

Since L_z does not appear in H, we are assured that $R(r)$ has no m dependence, so we only put E and ℓ in the subscript. Inserting this form into the above, we obtain

$$\left\{-\frac{\hbar^2}{2\,\mu}\left[\frac{1}{r^2}\,\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)-\frac{\ell(\ell+1)}{r^2}\right]+V(r)\right\}R_{E,\ell}(r) = E\,R_{E,\ell}(r)$$

This is the radial equation for a spherically symmetric potential in three dimensions.

Rotationally Invariant Problems in Three Dimensions (cont.)

We note for later reference that we have effectively factored the Hilbert space. The Hilbert space is

$$\mathbb{V} = \mathbb{V}_r \otimes \mathbb{V}_{\theta,\phi} = \mathbb{V}_r \otimes \left(\mathbb{V}_{\theta,\phi}^{(\ell=0)} \oplus \mathbb{V}_{\theta,\phi}^{(\ell=1)} \oplus \mathbb{V}_{\theta,\phi}^{(\ell=2)} \oplus \cdots \right)$$

Our eigenstates are of the form

$$|E,\ell,m\rangle = |E,\ell\rangle^{(r)} \otimes |\ell,m\rangle^{(\theta,\phi)}$$

$$\text{with} \qquad {}^{(r)}\langle r\,|E,\ell\rangle^{(r)} = R_{E,\ell}(r) \qquad \text{and} \qquad {}^{(\theta,\phi)}\langle\theta,\phi\,|\ell,m\rangle^{(\theta,\phi)} = Y_\ell^m(\theta,\phi)$$

We will drop the superscripts $^{(r)}$ and $^{(\theta,\phi)}$ on the states since there will not be ambiguity. We will keep them on the operators.

Our Hamiltonian has the form

$$H = \left[T^{(r)} + V^{(r)}\right] \otimes I^{(\theta,\phi)} + \frac{1}{2\,\mu}\left[R^{(r)}\right]^{-2} \otimes \left[L^2\right]^{(\theta,\phi)}$$

with

$$\begin{aligned}
\langle r,\theta,\phi\,|\left(\left[T^{(r)} + V^{(r)}\right] \otimes I^{(\theta,\phi)}\right)|E,\ell,m\rangle &= \langle r\,|\left[T^{(r)} + V^{(r)}\right]|E,\ell\rangle\langle\theta,\phi\,|\ell,m\rangle \\
\langle r\,|T^{(r)}|E,\ell\rangle &= -\frac{\hbar^2}{2\,\mu}\left[\frac{1}{r^2}\,\frac{\partial}{\partial r}\left(r^2\,\frac{\partial}{\partial r}\right)\right]R_{E,\ell}(r) \\
\langle r\,|V^{(r)}|E,\ell\rangle &= V(r)R_{E,\ell}(r) \\
\frac{1}{2\,\mu}\langle r,\theta,\phi\,|\left(\left[R^{(r)}\right]^{-2} \otimes \left[L^2\right]^{(\theta,\phi)}\right)|E,\ell,m\rangle &= \frac{1}{2\,\mu}\langle r\,|\left[R^{(r)}\right]^{-2}|E,\ell\rangle\langle\theta,\phi\,|L^2|\ell,m\rangle \\
&= \frac{1}{2\,\mu}\langle r\,|r^{-2}|E,\ell\rangle\langle\theta,\phi\,|\hbar^2\ell\,(\ell+1)\,|\ell,m\rangle \\
&= \frac{1}{2\,\mu}\,\frac{\hbar^2\,\ell\,(\ell+1)}{r^2}R_{E,\ell}(r)\,Y_\ell^m(\theta,\phi)
\end{aligned}$$

Generic Properties of Solutions of the Radial Equation

Simplifying the Radial Equation

What can we learn about solutions of the radial equation without detailed knowledge of the potential? We shall answer that question in this section.

We begin by rewriting the radial wavefunction as

$$R_{E,\ell}(r) = \frac{U_{E,\ell}(r)}{r}$$

because the radial equation then simplifies to the reduced radial equation

$$D_\ell(r)\, U_{E,\ell}(r) \equiv \left[-\frac{\hbar^2}{2\,\mu}\,\frac{d^2}{dr^2} + V(r) + \frac{\ell\,(\ell+1)\,\hbar^2}{2\,\mu\,r^2} \right] U_{E,\ell}(r) = E\, U_{E,\ell}(r)$$

This looks like the eigenvector-eigenvalue equation of the Hamiltonian in one dimension for a potential

$$V_{\mathit{eff}}(r) = V(r) + \frac{\ell\,(\ell+1)\,\hbar^2}{2\,\mu\,r^2}$$

with the additional restriction that r lies in the interval $[0, \infty)$. This effective potential, which includes a centrifugal barrier, will be familiar to those who have studied spherically symmetric potentials in classical mechanics.

Hermiticity Requirement

Since we now have something that looks like the eigenvector-eigenvalue equation for a one-dimensional Hamiltonian defined over an interval, and we have the Hermitian operator $D_\ell(r)$, we must check that the standard Hermiticity boundary condition is satisfied. That is, we require

$$\int_0^\infty dr\, U_1^* \left(D_\ell U_2\right) = \left[\int_0^\infty dr\, U_2^* \left(D_\ell U_1\right)\right]^*$$

The piece of this due to $V_{eff}(r)$ trivially satisfies the above because V_{eff} is a real numerical function, not an operator. So we must consider the derivative operator piece, which is

$$\int_0^\infty dr\, U_1^* \frac{d^2}{dr^2} U_2 = \int_0^\infty dr\, U_2 \frac{d^2}{dr^2} U_1^*$$

Generic Properties of Solutions of the Radial Equation (cont.)

Let's manipulate the RHS:

$$RHS = \int_0^\infty dr\, U_2 \frac{d^2}{dr^2} U_1^* = U_2 \frac{d}{dr} U_1^* \Big|_0^\infty - \int_0^\infty dr \frac{dU_2}{dr} \frac{dU_1^*}{dr}$$
$$= U_2 \frac{dU_1^*}{dr} \Big|_0^\infty - \frac{dU_2}{dr} U_1^* \Big|_0^\infty + \int_0^\infty dr \frac{d^2 U_2}{dr^2} U_1^*$$

So, to obtain $LHS = RHS$, we require

$$\left[U_2 \frac{dU_1^*}{dr} - \frac{dU_2}{dr} U_1^* \right] \Big|_0^\infty = 0$$
$$\frac{d}{dr} [U_2\, U_1^*] \Big|_0^\infty = 0$$

To evaluate the ∞ limit of the above, we need to know the behavior of $U_{E,\ell}(r)$ for $r \to \infty$. We can set this by requiring normalizability. We already know the angular part of the wavefunction is normalized to unity when the solid angle integral is done. The radial portion of the normalization integral is

$$\int_0^\infty dr\, r^2 |R_{E,\ell}(r)|^2 = \int_0^\infty dr\, |U_{E,\ell}(r)|^2$$

In order for this to equal unity, we need $U_{E,\ell}(r) \to 0$ as $r \to \infty$. To make it normalizable to a delta function, we require $U_{E,\ell}(r) \to e^{i\,k\,r}$ as $r \to \infty$ just as we would for any one-dimensional problem. For the decaying case, the $r \to \infty$ term vanishes. For the oscillating case, just as we showed in the case of plane waves in one dimension, the upper limit also vanishes. So, we are left with the condition

$$\frac{d}{dr}\left[U_2\, U_1^*\right]\Big|_0 = 0$$

Clearly, the function acted on by the derivative must converge to a constant as $r \to 0$. Since U_1 and U_2 are arbitrary, each one must converge to a constant separately. So we have the additional requirement

$$U(r) \stackrel{r\to 0}{\longrightarrow} c$$

with c a constant.

Form of Solutions for $r \rightarrow 0$

We have established that, in order for the $D_\ell(r)$ operator to be Hermitian, we require $U_{E,\ell}(r) \rightarrow c$ as $r \rightarrow 0$ and $U_{E,\ell}(r) \rightarrow 0$ or $e^{i\,k\,r}$ as $r \rightarrow \infty$. Now let us check whether these requirements are consistent with the eigenvector-eigenvalue equation.

Let's first check $r \rightarrow 0$. It is insufficient to check the one-dimensional equation for $U_{E,\ell}(r)$ because the relation $R_{E,\ell}(r) = U_{E,\ell}(r)/r$ breaks down at $r \rightarrow 0$ unless $U_{E,\ell}(r) \rightarrow 0$ as fast as r. So we need to check the eigenvector-eigenvalue equation for the full Hamiltonian of a particle in three dimensions. That is

$$-\frac{\hbar^2}{2\,\mu}\,\nabla^2\psi_{E,\ell,m}(r,\theta\,\phi) + V(r)\,\psi_{E,\ell,m}(r,\theta,\phi) = E\,\psi_{E,\ell,m}(r,\theta,\phi)$$

which, based on our asymptotic form for $r \rightarrow 0$, reduces to

$$\left[-\frac{\hbar^2}{2\,\mu}\,\nabla^2 + V(r) + \frac{\hbar^2\ell\,(\ell+1)}{2\,\mu\,r^2}\right]\frac{1}{r} = E\,\frac{1}{r}$$

Generic Properties of Solutions of the Radial Equation (cont.)

This is problematic in two ways. First, for $r \neq 0$, $\nabla^2(1/r) = 0$ for $r \neq 0$, so we obtain

$$V(r) \stackrel{r \to 0}{\longrightarrow} E - \frac{\hbar^2 \ell\,(\ell+1)}{2\,\mu\,r^2}$$

This is not a very generic form for the potential near the origin.

The other problem occurs at the origin. The derivatives involved in $\nabla^2(1/r)$ become infinite as $r \to 0$. We have to resort to Gauss's Law to determine its value in a rigorous manner. Consider the integral of $\nabla^2(1/r)$ over the sphere of radius r, whose volume we will write as $\mathcal{V}(r)$ and whose surface is $\mathcal{S}(r)$. We can transform the integral using Gauss's Law:

$$\begin{aligned}\int_{\mathcal{V}(r)} d\Omega\,(r')^2\,dr'\,\nabla^2\frac{1}{r'} &= \int_{\mathcal{S}(r)} d\Omega\,r^2\,\hat{r}\cdot\vec{\nabla}\frac{1}{r'} \\ &= 4\,\pi r^2\,\frac{d}{dr}\,\frac{1}{r} \\ &= -4\,\pi\end{aligned}$$

The integral is independent of the size of the volume and yields a constant. This is exactly the characteristic of a delta function. So we conclude

$$\nabla^2 \frac{1}{r} = 4\,\pi\,\delta(\vec{r}) \equiv 4\,\pi\,\delta(x)\,\delta(y)\,\delta(z)$$

Plugging this back into the eigenvector-eigenvalue equation, we see the it can only be satisfied if $V(r) \longrightarrow \delta(\vec{r})$ at the origin. Again, a very special case that we will in general not be interested in.

The only way to resolve the above problems is to set $c = 0$. That is, for $r \longrightarrow 0$, $U_{E,\ell}(r) \longrightarrow 0$. How quickly $U_{E,\ell}(r)$ must vanish will be studied next.

Generic Properties of Solutions of the Radial Equation (cont.)

Let's now assume that $V(r)$ is less singular than $1/r^2$ for $r \to 0$. The reasons for doing so are: 1) most physical potentials, such as the gravitational and Coulomb potential, satisfy this condition; and 2) this allows us to assume the centrifugal term dominates near the origin for $\ell \neq 0$, so the exact form of the potential becomes unimportant there and we may derive generic properties. In this limit, the one-dimensional equation reduces to

$$\frac{d^2}{dr^2} U_\ell(r) = \frac{\ell(\ell+1)}{r^2} U_\ell(r)$$

where we have dropped the E term and hence the E subscript because the E term becomes negligible as the centrifugal barrier dominates for $r \to 0$. The appearance of two powers of r in the denominator when two derivatives are taken suggests power law behavior; assuming $U_\ell(r) = r^\gamma$ implies

$$\gamma(\gamma-1) = \ell(\ell+1) \qquad \Longrightarrow \qquad \gamma = \ell+1 \quad \text{or} \quad \gamma = -\ell$$

also known as the *regular* and *irregular* solutions because of their behavior near the origin. The latter one fails our condition $U_{E,\ell}(r) \to 0$ for $r \to 0$, so we keep only the regular solution. Since $\ell \geq 1$, we are assured that $R_{E,\ell}(r) = U_{E,\ell}(r)/r \to 0$ as $r \to 0$. Hence, there is no probability for finding the particle at the origin, which is consistent with the infinitely large potential barrier there.

Generic Properties of Solutions of the Radial Equation (cont.)

For $\ell = 0$, which we did not consider above, the form of the solution will now depend on the potential, and possibly also on E if the potential goes to a constant or vanishes at the origin. Nothing generic can be said. In the context of the hydrogen atom, one can see that the above form $U_\ell(r) = r^{\ell+1}$ is also valid for $\ell = 0$ for the Coulomb potential. This results in $R_{E,\ell}(r) \longrightarrow r/r = 1$ as $r \longrightarrow 0$, so the absence of a centrifugal barrier allows the particle to be found at the origin.

Form of Solutions for $r \to \infty$

Let's consider potentials for which $V(r) \to 0$ as $r \to \infty$ so that the form of the potential in this limit is unimportant. There are important violations of this condition — the SHO, for example — that must be considered case-by-case. But, for potentials that vanish at ∞, the eigenvector-eigenvalue equation reduces to

$$\frac{d^2}{dr^2} U_E(r) = -\frac{2\,\mu\,E}{\hbar^2}\, U_E(r)$$

We see that the dependence on ℓ vanishes because the centrifugal barrier term becomes negligible, so we label solutions by E alone. The form of the solution to the above equation is an exponential, though whether it has real or imaginary argument depends on the sign of E.

Generic Properties of Solutions of the Radial Equation (cont.)

For $E > 0$, the exponential has an imaginary argument and our solutions are of the form

$$U_E(r) = A\, e^{i\,k\,r} + B\, e^{-i\,k\,r} \qquad k = \frac{1}{\hbar}\sqrt{2\,\mu\,E}$$

which makes sense, since it looks like a free particle.

Let's consider how this asymptotic form must match onto the form at smaller r. Since the solution must have $r^{\ell+1}$ dependence near the origin, and the exponential form cannot provide this, there must be a multiplying term that matches onto the power-law dependence at small r and that becomes constant at large r. To determine this factor, we want to consider the problem in the regime where $V(r)$ is not negligible but is small compared to E: we will thus see the multiplying factor transition from a power law to a constant. It should be clear that this regime is perfect for applying the WKB approximation: for $V(r) = 0$ exactly, the wavefunction's wavelength will be constant, $\lambda = 2\,\pi/k$, but for $0 \neq |V(r)/E| \ll 1$, λ will be slowly varying.

Generic Properties of Solutions of the Radial Equation (cont.)

That is, we start with the standard WKB form

$$U_E(r) = e^{\pm \frac{i}{\hbar}\,\phi(r)} \qquad \phi(r) = \int_{r_0}^{r} dr\,\sqrt{2\,\mu\,(E - V(r))}$$

(r_0 is some arbitrary reference point at large enough r that $|V(r)/E| \ll 1$ holds)
Making the approximation $|V(r)/E| \ll 1$ lets us Taylor expand the square root:

$$\begin{aligned}\phi(r) &= \int_{r_0}^{r} dr'\,\sqrt{2\,\mu\,E}\left(1 - \frac{1}{2}\,\frac{V(r')}{E}\right) = \hbar\,k\int_{r_0}^{r} dr' - \frac{\hbar\,k}{2\,E}\int_{r_0}^{r} dr'\,V(r') \\ &= \hbar\,k\,(r - r_0) - \frac{\mu}{\hbar\,k}\int_{r_0}^{r} dr'\,V(r')\end{aligned}$$

So we have

$$U_E(r) = f(r_0)\,e^{\pm i\,k\,r}\,\exp\left(\mp\frac{i}{\hbar}\,\frac{\mu}{\hbar\,k}\int_{r_0}^{r} dr'\,V(r')\right)$$

where $f(r_0)$ is a normalization factor that depends on the choice of lower limit of the integral.

Generic Properties of Solutions of the Radial Equation (cont.)

Now, we want the $V(r)$ integral to converge as $r \to \infty$ so we recover the pure plane-wave behavior; that is, we need

$$\int_{r_0}^{\infty} dr' V(r') = c(r_0)$$

In order for the integral to converge, we need $V(r)$ to fall off faster than $1/r$; $V(r) = 1/r$ will make the value of the integral depend logarithmically on the infinite upper limit, yielding an infinite integral. This condition may be rewritten as

$$r\, V(r) \overset{r\to\infty}{\longrightarrow} 0$$

(When we consider scattering, the lack of convergence of the integral for potentials that fall off like $1/r$ or slower will manifest as an infinite cross-section for scattering.)

Note that, because there are two allowed solutions (the $\pm$ signs), there are two coefficient degrees of freedom. These will be determined by requiring this solution to match onto the wavefunction for smaller r. Because there are two matching conditions (wavefunction and its derivative) and two degrees of freedom, we expect no restriction on k and hence no quantization of energies for $E > 0$.

Generic Properties of Solutions of the Radial Equation (cont.)

For $E < 0$, the exponential has real argument. As usual, we must reject the growing solution because it cannot be normalized, so we have

$$U_E(r) \to e^{-\kappa r} \qquad \kappa = \frac{1}{\hbar}\sqrt{2\mu\,|E|}$$

We may repeat the WKB procedure above to find

$$U_E(r) = f(r_0)\, e^{-\kappa r}\, \exp\left(-\frac{1}{\hbar}\,\frac{\mu}{\hbar\kappa}\int_{r_0}^{r} dr'\, V(r')\right)$$

which yields the same condition on $V(r)$ to yield the purely decaying form for $r \to \infty$. We note, though, that for the Coulomb potential, the logarithmic dependence of the integral on the upper limit manifests as a power law in r:

$$\int_{r_0}^{r} dr'\, V(r') = -e^2 \int_{r_0}^{r} dr'\, \frac{1}{r'} = -e^2\,(\ln r - \ln r_0)$$

$$\Longrightarrow \qquad U_E(r) = g(r_0)\, r^{\,\mu\, e^2/\hbar^2\kappa}\, e^{-\kappa r}$$

(In the free-particle case, the logarithm is multiplied by i, so it does not result in a power law in r, but in a nonconvergent complex phase.)

Generic Properties of Solutions of the Radial Equation (cont.)

κ, and hence E, will become discretized by the requirement that $U_E(r)$ match onto a solution valid for moderate and small r: there will be two matching conditions (wavefunction and derivative) but only one normalization degree of freedom to use, so the freedom in κ will be reduced by the other condition. This is the standard manner of obtaining discretization of bound states in one dimension.

One can show that the eigenfunctions of D_ℓ with $E < 0$ and $U \stackrel{r \to 0}{\longrightarrow} 0$ are nondegenerate in the same way that we showed that bound states of the Hamiltonian for one-dimensional systems are nondegenerate.

Hence, we are assured that there is an orthonormalization condition for bound states

$$\int_0^\infty dr\, U_{E,\ell}(r)\, U_{E',\ell}(r) = \delta_{E,E'}$$

or, using $\psi_{E,\ell,m}(r) = R_{E,\ell}(r)\, Y_\ell^m(\theta,\phi)$, the orthonormalization condition for the full 3-dimension bound eigenstates is

$$\int_0^\infty r^2\, dr \int_0^\pi s_\theta\, d\theta \int_0^{2\,\pi} d\phi\; \psi^*_{E,\ell,m}(r,\theta,\phi)\, \psi_{E',\ell',m'}(r,\theta,\phi) = \delta_{E,E'}\delta_{\ell,\ell'}\delta_{m,m'}$$

Of course, the bound states are orthogonal to the free states, and the free states will satisfy a delta-function normalization in E that we will discuss below.

The Free Particle

This example is discussed in detail in Shankar, so we only quote the results here. It is instructive to read Shankar, though, because the technique used to obtain the radial solutions is an interesting application of raising and lowering operators. One could just do the standard series solution for the differential equation, too.

For the free particle, the reduced radial equation becomes

$$\left[\frac{d^2}{dr^2} + k^2 - \frac{\ell\,(\ell+1)}{r^2}\right] U_{E,\ell}(r) = 0 \qquad k = \frac{1}{\hbar}\sqrt{2\,\mu\,E}$$

One can solve the equation easily for $\ell = 0$, and the equation looks like the SHO problem with r^2 replaced by $1/r^2$, so one is prompted to define raising and lowering operators, solve the $\ell = 0$ case, directly, and use the raising operators to obtain the $\ell \neq 0$ solutions.

Solutions for Specific Rotationally Invariant Potentials (cont.)

One can also solve the problem using series solution techniques, though there will be no termination requirement since the solutions should converge to $e^{\pm i\,k\,r}$. The resulting solutions for $R(r)$ are called the spherical Bessel functions, $j_n(\rho)$ and spherical Neumann functions, $\eta_n(\rho)$, where $\rho = k\,r$. The first two of each of these are

$$j_0(\rho) = \frac{\sin\rho}{\rho} \qquad \eta_0(\rho) = -\frac{\cos\rho}{\rho}$$

$$j_1(\rho) = \frac{\sin\rho}{\rho^2} - \frac{\cos\rho}{\rho} \qquad \eta_1(\rho) = -\frac{\cos\rho}{\rho^2} - \frac{\sin\rho}{\rho}$$

These functions have asymptotic forms

$$j_\ell(\rho) \stackrel{\rho\to 0}{\longrightarrow} \frac{\rho^\ell}{(2\,\ell+1)!!} \qquad \eta_\ell(\rho) \stackrel{\rho\to 0}{\longrightarrow} -\frac{(2\,\ell-1)!!}{\rho^{\ell+1}}$$

$$j_\ell(\rho) \stackrel{\rho\to\infty}{\longrightarrow} \frac{1}{\rho}\,\sin\left(\rho - \ell\,\frac{\pi}{2}\right) \qquad \eta_\ell(\rho) \stackrel{\rho\to\infty}{\longrightarrow} -\frac{1}{\rho}\,\cos\left(\rho - \ell\,\frac{\pi}{2}\right)$$

Solutions for Specific Rotationally Invariant Potentials (cont.)

Only the spherical Bessel functions are regular as $\rho \longrightarrow 0$, so they are the solutions we are allowed to use. This reflects the matching conditions to the solution at the origin and discards either the cos-like or sin-like solution for each value of ℓ. The full solutions are then

$$\begin{aligned}\psi_{E,\ell,m}(r,\theta,\phi) &= \langle r,\theta\,\phi\,|E,\ell,m\,\rangle \\ &= \left(\langle r\,|\otimes\langle\theta,\phi\,|\right)\left(|E,\ell\,\rangle\otimes|\ell,m\,\rangle\right) \\ &= \langle r\,|E,\ell\,\rangle\langle\theta,\phi\,|\ell,m\,\rangle \\ &= j_\ell(k\,r)\,Y_\ell^m(\theta,\phi) \qquad k=\frac{1}{\hbar}\sqrt{2\,\mu\,E}\end{aligned}$$

where we have written the solution as a direct product of states in the $\mathbb{V}_r$ Hilbert space that describes the radial behavior and the $\mathbb{V}_{\theta,\phi}$ Hilbert space that describes the angular behavior. The orthonormalization relation is

$$\begin{aligned}&\int_0^\infty r^2 dr \int_0^\pi s_\theta\, d\theta \int_0^{2\,\pi} d\phi\; \psi^*_{E,\ell,m}(r,\theta,\phi)\,\psi_{E',\ell',m'}(r,\theta,\phi) \\ &= \frac{2}{\pi\,k^2}\,\delta(k-k')\delta_{\ell,\ell'}\delta_{m,m'}\end{aligned}$$

Relation to Free Particle Solutions in Cartesian Coordinates

Again, we don't repeat in detail what is in Shankar. If we consider the special case of a free particle state of well-defined momentum $\vec{p} = p\widehat{z}$, the wavefunction is

$$\langle x, y, z \,|\vec{p} = p\widehat{z}\,\rangle = \left(\frac{1}{2\,\pi}\right)^{3/2} e^{\frac{i}{\hbar}\,p\,z} \quad \Longleftrightarrow \quad \langle r, \theta, \phi \,|\vec{p} = p\widehat{z}\,\rangle = \left(\frac{1}{2\,\pi}\right)^{3/2} e^{\frac{i}{\hbar}\,p\,r\,c_\theta}$$

One can show (see Shankar for details)

$$\begin{aligned} e^{i\,k\,r\,c_\theta} &= \sum_{\ell=0}^{\infty} i^\ell\,(2\ell+1)\,j_\ell(k\,r)P_\ell(c_\theta) \\ \Longleftrightarrow \qquad |\vec{p} = p\widehat{z}\,\rangle &= \sum_{\ell=0}^{\infty} i^\ell\,(2\ell+1)\left|\, E = \frac{p^2}{2\,\mu}, \ell, m=0\right\rangle \\ &= \sum_{\ell=0}^{\infty} i^\ell\,(2\ell+1)\left|\, E = \frac{p^2}{2\,\mu}, \ell\right\rangle \otimes |\ell, m=0\,\rangle \end{aligned}$$

where we have written the eigenstate $|E, \ell, m\,\rangle$ of the Hamiltonian in factorized form as explained earlier.

Solutions for Specific Rotationally Invariant Potentials (cont.)

Then, one can obtain a state with the momentum directed in an arbitrary direction $\vec{p}$ by applying the appropriate rotation operator to the above state. That is, suppose one wants to know the free particle state with momentum operator eigenvalue

$$\vec{p} = \widehat{x}\, p\, s_{\theta_p} c_{\phi_p} + \widehat{y}\, p\, s_{\theta_p} s_{\phi_p} + \widehat{z}\, p\, c_{\theta_p}$$

The classical vector $\vec{p}$ is obtained from $p\widehat{z}$ by a rotation by an angle θ_p around the direction $\widehat{x}s_{\phi_p} - \widehat{y}c_{\phi_p}$ (you can check this explicitly). Thus, the QM state $|\vec{p}\,\rangle$ is obtained from $|p\widehat{z}\,\rangle$ by the corresponding QM rotation operator:

$$\begin{aligned}
|\vec{p}\,\rangle &= T\left(\theta_p \left[\widehat{x}s_{\phi_p} - \widehat{y}c_{\phi_p}\right]\right) |\vec{p} = p\widehat{z}\,\rangle \\
&= \sum_{\ell=0}^{\infty} i^{\ell}\,(2\ell+1) \left| E = \frac{p^2}{2\,\mu}, \ell \right\rangle \otimes T^{(\ell)}\left(\theta_p \left[\widehat{x}s_{\phi_p} - \widehat{y}c_{\phi_p}x\right]\right) |\ell, m = 0\,\rangle
\end{aligned}$$

where $T^{(\ell)}$ is the identity operator in $\mathbb{V}_r$, the space in which $|E, \ell\,\rangle$ lives, and acts on the $|\ell, m\,\rangle$ portion of the state in the manner that we described earlier for the action of rotation operators in the $|j, m\,\rangle$ basis.

Section 15

Spin Angular Momentum

Lecture 50:
Spin in Quantum Mechanics
Review of Classical Cartesian Tensors

Date Revised: 2009/03/07
Date Given: 2009/02/25

Formulation

So far, we have considered particles whose position-basis representations consist of a single number at every point in space — this is the position-basis wavefunction. However, it is found empirically that most fundamental particles and many composite particles have associated with them an orientation information that is not encoded in the spatial wavefunction. This orientation information seems to act empirically like an angular momentum.

The archetypal examples of this effect are the Stern-Gerlach experiment and the precession of a charged particle with spin in a magnetic field. In the former, it is observed that, in addition to the spatial wavefunction, electrons and atoms can have an additional degree of freedom that looks like a magnetic dipole moment, and a magnetic field can be used to separate particles that have this moment aligned or anti-aligned with the magnetic field. In the latter, it is observed that this magnetic dipole momentum appears to be associated with an angular momentum, as if the charged particle were a spinning ball of charge.

For these empirical reasons, we are led to ask whether there is a way to formulate the concept of not only a position-dependent wavefunction, but one that carries some sort of "spin" information at each point in space. To formulate such a concept, we need to return to the concepts of vectors and tensors in classical mechanics; we will use them to define a concept of spin in QM.

Review of Cartesian Tensors in Classical Mechanics

Prologue

We have so far relied on your intuitive grasp of what a vector is and how it is affected by a classical rotation. We need to formalize this intuition into a proper definition of scalars, vectors, and tensors, Cartesian and spherical, so we may define a sensible extension in quantum mechanics.

Cartesian Tensors — Definitions

In Section 14.2, we discussed passive and active coordinate transformations of classical vectors. In discussing tensors, we will be interested in passive transformations because we will define tensors in terms of the way the coordinate representations of a given tensor in two different coordinate systems are related.

Recall that a vector $\vec{a}$ has two different representations (a_x, a_y, a_z) and $(a_{x'}, a_{y'}, a_{z'})$ in two different coordinate systems F and F' with coordinate axes (x, y, z) and (x', y', z'), with the latter representation obtained from the former by application of $R_{-\vec{\theta}}$ where $\vec{\theta}$ indicates how F' is rotated relative to F:

$$\begin{bmatrix} a_{x'} \\ a_{y'} \\ a_{z'} \end{bmatrix} = \begin{bmatrix} c_\theta & s_\theta & 0 \\ -s_\theta & c_\theta & 0 \\ 0 & & 1 \end{bmatrix} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} = \mathbf{R}_{P,\theta\hat{z}} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}$$

$$\text{with} \qquad \mathbf{R}_{P,\theta\hat{z}} = \mathbf{R}_{A,-\theta\hat{z}} = \mathbf{R}_{-\theta\hat{z}}$$

Review of Cartesian Tensors in Classical Mechanics (cont.)

Here, we generalize this idea to an object called a tensor that, in a similar way, has different coordinate representations in coordinate systems that are related to each other by rotation, but itself has an abstract existence independent of and unchanged by the choice of coordinate system.

Specifically, if we consider two sets of coordinates axes F and F' that are related to each other by a rotation, then the coordinate representation of a tensor in F consists of a set of numbers (how many depends on the *rank* of the tensor and whether it is *Cartesian* or *spherical*, which we will discuss below), and the tensor is *defined* by the fact that its representation in F' is related to its representation in F by a specific set of transformation laws involving the rotation matrices we discussed in Section 14.2 and 14.6. Cartesian tensors are transformed using the rotation matrices generated by the $\vec{\mathbf{M}}$ of Section 14.2 (the $\mathbf{R}_{-\vec{\theta}}$ shown above), while spherical tensors are transformed using the rotation matrices generated by the $|j, m\rangle$ basis representation of $\vec{J}$ found in Section 14.6, the matrix representations of the $T^{(j)}(\vec{\theta})$.

Review of Cartesian Tensors in Classical Mechanics (cont.)

You are certainly familiar with two kinds of tensors. The first, called a scalar, is also known as a rank 0 tensor (Cartesian or spherical). A scalar is essentially a trivial tensor because its transformation law is that its representation in any coordinate system is a single number and that this number is the same in any two coordinate systems related by a rotation. Examples include the mass of a particle, the total energy of a particle, etc.

The second kind of tensor you are familiar with is called a vector or rank 1 Cartesian tensor. As you know, the coordinate representation of a vector $\vec{v}$ in a particular coordinate system F consists of N numbers (N is the number of spatial dimensions, $N = 3$ for what are considering), which we shall denote by $\underline{(\vec{v})}$ or $\{\underline{(\vec{v})}_j\}$. Its representation in a different frame F', which we shall denote by $\underline{(\vec{v})}'$ or $\{\underline{(\vec{v})}'_j\}$, is related to that in F by

$$\underline{(\vec{v})}' = \mathbf{R}_{-\vec{\theta}}\,\underline{(\vec{v})} \tag{15.1}$$

$$\underline{(\vec{v})}'_j = \sum_k \left[\mathbf{R}_{-\vec{\theta}}\right]_{jk} \underline{(\vec{v})}_k$$

(We need the underscore to distinguish $\vec{v}$ from its coordinate representation $\underline{(\vec{v})}$ and we need the parentheses later on when we might consider a different vector $\vec{v}\,'$ and need to make the distinction between $\underline{(\vec{v})}'$, the coordinate representation $\vec{v}$ in F', and $\underline{(\vec{v}\,')}$, the coordinate representation $\vec{v}\,'$ in F.)

The eabove equation is the same as the matrix equation we wrote down just a couple of slides back, except that we have changed our convention on how primes are used for the sake of brevity of notation: rather than putting the prime on the subscript as we did earlier to indicate different coordinate representations of the same vector, we are putting the prime on the vector representation or components themselves.

Let's consider the relations between the unit vectors of the F and F' frames via the rotation matrix to be sure these relationships are clear. Denote the unit vectors of the F frame by $\{\vec{e}_j\}$ and those of the F' frame by $\{\vec{e}_j'\}$ where j runs from 1 to 3 for three spatial dimensions. The $\{\vec{e}_j\}$ are what we earlier would have called $\widehat{x}$, $\widehat{y}$, and $\widehat{z}$. The $\{\vec{e}_j'\}$ are what we would have called $\widehat{x}'$, $\widehat{y}'$, and $\widehat{z}'$. Since we have written the prime on a vector, not a coordinate representation, the prime is telling us $\vec{e}_j'$ is a different vector than $\vec{e}_j$.

Here are the coordinate representations in their natural frames and how they are related:

$$\underline{(\vec{e}_1)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \underline{(\vec{e}_1')}' \qquad \underline{(\vec{e}_2)} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \underline{(\vec{e}_2')}' \qquad \underline{(\vec{e}_3)} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \underline{(\vec{e}_3')}'$$

The left side of each equation is the coordinate representation of an $\vec{e}_j$ in the F frame (hence the lack of primes) while the right side is the coordinate representation of an $\vec{e}_j'$ in the F' frame; hence the prime both inside the parentheses, referring to the vector, and outside the parentheses, referring to the frame of the coordinate representation. The above simply states that the $\{\vec{e}_j\}$ are the unit vectors of the F frame and the $\{\vec{e}_j'\}$ are the unit vectors of the F' frame and that they, respectively, correspond to the basis elements of the F and F' coordinate representations. You can see that coordinate representations and matrix representations are not very different from one another; the former is just a special case of the latter for a Hilbert space consisting of vectors in three spatial dimensions.

Now, let's see how the rotation matrices relate the coordinate representations between frames. Again, F' is obtained from F by rotation by an angle θ CCW about the $\widehat{\theta}$ direction. Then the coordinate representations in the F' frame of the F frame unit vectors $\{\vec{e}_j\}$ are obtained from their coordinate representation sin the F frame by

$$\underline{(\vec{e}_i)'} = \mathbf{R}_{-\vec{\theta}}\,\underline{(\vec{e}_i)}$$

$$\underline{(\vec{e}_1)'} = \mathbf{R}_{-\vec{\theta}}\begin{bmatrix}1\\0\\0\end{bmatrix} \qquad \underline{(\vec{e}_2)'} = \mathbf{R}_{-\vec{\theta}}\begin{bmatrix}0\\1\\0\end{bmatrix} \qquad \underline{(\vec{e}_3)'} = \mathbf{R}_{-\vec{\theta}}\begin{bmatrix}0\\0\\1\end{bmatrix}$$

This is just a special case of our generic transformation rule for the coordinate representations of a vector, Equation 15.1. Remember, we use $\mathbf{R}_{-\vec{\theta}}$ because we are not actively rotating the vectors $\{\vec{e}_j\}$, we are writing their representations in a different frame F'.

Similarly, the coordinate representations of the F' frame unit vectors $\{\vec{e}_j{}'\}$ in the F' frame are obtained from those in the F frame by

$$\underline{\left(\vec{e}_j{}'\right)'} = \mathbf{R}_{-\vec{\theta}}\underline{\left(\vec{e}_j{}'\right)}$$

However, this is not usually what one wants because, in this case, we know the $\left\{\underline{\left(\vec{e}_j{}'\right)'}\right\}$ are simple and the $\left\{\underline{\left(\vec{e}_j{}'\right)}\right\}$ are not. Rather, we want the inverse equation:

$$\underline{\left(\vec{e}_j{}'\right)} = \mathbf{R}_{\vec{\theta}}\underline{\left(\vec{e}_j{}'\right)'}$$

$$\underline{\left(\vec{e}_1{}'\right)} = \mathbf{R}_{\vec{\theta}}\begin{bmatrix}1\\0\\0\end{bmatrix} \qquad \underline{\left(\vec{e}_2{}'\right)} = \mathbf{R}_{\vec{\theta}}\begin{bmatrix}0\\1\\0\end{bmatrix} \qquad \underline{\left(\vec{e}_3{}'\right)} = \mathbf{R}_{\vec{\theta}}\begin{bmatrix}0\\0\\1\end{bmatrix}$$

where now $\mathbf{R}_{\vec{\theta}}$ is used because we are obtaining the F frame coordinate representation from the F' frame representations and the F frame is obtained from the F' frame by rotation by $-\theta$ about $\widehat{\theta}$.

Now, let's generalize. A rank n Cartesian tensor is an object that satisfies similar relations between coordinate representations, but involving more rotation matrices and more numbers in each representation. Specifically, a rank n Cartesian tensor $\mathcal{T}$ is an object that has coordinate representation $\underline{(\mathcal{T})}$ with N^n components $\underline{(\mathcal{T})}_{j_1\cdots j_n}$ (where N is the dimensionality of the physical space, $N = 3$ for what we are considering) with passive transformation properties

$$\underline{(\mathcal{T})}'_{j_1\cdots j_n} = \sum_{k_1,k_2,\cdots,k_n} \left[\mathbf{R}_{-\vec{\theta}}\right]_{j_1 k_1} \cdots \left[\mathbf{R}_{-\vec{\theta}}\right]_{j_n k_n} \underline{(\mathcal{T})}_{k_1\cdots k_n} \tag{15.2}$$

We see why a scalar is a rank 0 tensor and a vector is a rank 1 tensor. We will in general use the *Einstein Summation Convention* to write the above as

$$\underline{(\mathcal{T})}'_{j_1\cdots j_n} = \left[\mathbf{R}_{-\vec{\theta}}\right]_{j_1 k_1} \cdots \left[\mathbf{R}_{-\vec{\theta}}\right]_{j_n k_n} \underline{(\mathcal{T})}_{k_1\cdots k_n}$$

Review of Cartesian Tensors in Classical Mechanics (cont.)

A rank 2 tensor has coordinate representations that look like square $N \times N$ matrices; what distinguishes a rank 2 tensor from a simple matrix is the relation between the coordinate representations in different frames. It is important to remember the distinction! However, this form lets us write the transformation law in a compact manner, like we do for vectors:

$$\underline{(\mathcal{T})}'_{jk} = \left[\mathbf{R}_{-\vec{\theta}}\right]_{jm} \left[\mathbf{R}_{-\vec{\theta}}\right]_{kn} \underline{(\mathcal{T})}_{mn} = \left[\mathbf{R}_{-\vec{\theta}}\right]_{jm} \underline{(\mathcal{T})}_{mn} \left[\mathbf{R}^{T}_{-\vec{\theta}}\right]_{nk} = \left[\mathbf{R}_{-\vec{\theta}}\right]_{jm} \underline{(\mathcal{T})}_{mn} \left[\mathbf{R}^{-1}_{-\vec{\theta}}\right]_{nk}$$

$$\underline{(\mathcal{T})}' = \mathbf{R}_{-\vec{\theta}}\, \underline{(\mathcal{T})}\, \mathbf{R}^{T}_{-\vec{\theta}} = \mathbf{R}\, \underline{(\mathcal{T})}\, \mathbf{R}^{-1}_{-\vec{\theta}}$$

where $\underline{(\mathcal{T})}$ and $\underline{(\mathcal{T})}'$ are $N \times N$ matrices and $\mathbf{R}^{T} = \mathbf{R}^{-1}$ follows from the fact that rotation matrices are orthogonal matrices (so $\mathbf{R}^{T}_{\vec{\theta}} = \mathbf{R}^{-1}_{\vec{\theta}}$). The last expression is the similarity transformation of the $N \times N$ matrix $\underline{(\mathcal{T})}$ by the orthogonal matrix $\mathbf{R}_{-\vec{\theta}}$.

Tensors — Examples

- The norm of a Cartesian vector is a scalar:

$$\underline{(\vec{v})}' \cdot \underline{(\vec{v})}' = \underline{(\vec{v})}'_j \, \underline{(\vec{v})}'_j = \left[\mathbf{R}_{-\vec{\theta}}\right]_{jm} \underline{(\vec{v})}_m \left[\mathbf{R}_{-\vec{\theta}}\right]_{jn} \underline{(\vec{v})}_n$$
$$= \delta_{mn} \, \underline{(\vec{v})}_m \, \underline{(\vec{v})}_n = \underline{(\vec{v})}_m \, \underline{(\vec{v})}_m = \underline{(\vec{v})} \cdot \underline{(\vec{v})}$$

  where we have used $\underline{(\vec{v})}' = \mathbf{R}_{-\vec{\theta}} \, \underline{(\vec{v})}$ and the orthonormality property $\left[\mathbf{R}_{-\vec{\theta}}\right]_{jm} \left[\mathbf{R}_{-\vec{\theta}}\right]_{jn} = \delta_{mn}$ (*i.e.*, $\mathbf{R}^T = \mathbf{R}^{-1}$). We see that the value of the norm of a Cartesian vector is the same in the two frames, hence the coordinate representations of the norm are identical and it is a scalar. The dot product of any two vectors is a scalar by a similar argument.

- An obvious rank 2 Cartesian tensor is the *outer product* of two vectors:

$$\underline{(\mathcal{T})}_{jk} = \underline{(\vec{a})}_j \, \underline{(\vec{b})}_k \qquad \text{or} \qquad \mathcal{T} = \vec{a} \, \vec{b}^{\,T}$$

  Since each Cartesian vector transforms as a rank 1 Cartesian tensor, it is obvious that the above product transforms as a rank 2 Cartesian tensor:

$$\underline{(\mathcal{T})}'_{jk} = \underline{(\vec{a})}'_j \underline{(\vec{b})}'_k = \left[\mathbf{R}_{-\vec{\theta}}\right]_{jm} \underline{(\vec{a})}_m \left[\mathbf{R}_{-\vec{\theta}}\right]_{kn} \underline{(\vec{b})}_n = \left[\mathbf{R}_{-\vec{\theta}}\right]_{jm} \left[\mathbf{R}_{-\vec{\theta}}\right]_{kn} \underline{(\mathcal{T})}_{mn}$$

- More generally, if we take a rank m Cartesian tensor $\mathcal{U}$ with coordinate representation components $\underline{(\mathcal{U})}_{j_1 \cdots j_m}$ and a rank n Cartesian tensor $\mathcal{V}$ with coordinate representation components $\underline{(\mathcal{V})}_{k_1 \cdots k_n}$ and **contract** over — *i.e.*, match up indices and sum, the generalization of a dot product — any p pairs of indices, then the resulting set of quantities is a rank $m + n - 2p$ Cartesian tensor. Proving it is clearly a tedious exercise in index arithmetic relying on the rotation matrix orthogonality relation $\left[\mathbf{R}_{-\vec{\theta}}\right]_{mj} \left[\mathbf{R}_{-\vec{\theta}}\right]_{mk} = \delta_{jk}$ and its transpose relation $\left[\mathbf{R}_{-\vec{\theta}}\right]_{jm} \left[\mathbf{R}_{-\vec{\theta}}\right]_{km} = \delta_{jk}$, much like the proof that the norm of a Cartesian vector is a scalar. Taking $p = 0$ as a special case gives us the simple outer product of the two Cartesian tensors, which reduces to the previous example when both Cartesian tensors are rank 1.

- The identity matrix is a rank 2 Cartesian tensor and, in fact, it is **isotropic**, meaning that its coordinate representation is the same in all frames. Let's just try transforming it to see this:

$$\begin{aligned}\underline{(\mathcal{I})}'_{jk} &= \left[\mathbf{R}_{-\vec{\theta}}\right]_{jm}\left[\mathbf{R}_{-\vec{\theta}}\right]_{kn}\underline{(\mathcal{I})}_{mn} = \left[\mathbf{R}_{-\vec{\theta}}\right]_{jm}\left[\mathbf{R}_{-\vec{\theta}}\right]_{kn}\delta_{mn} \\ &= \left[\mathbf{R}_{-\vec{\theta}}\right]_{jm}\left[\mathbf{R}_{-\vec{\theta}}\right]_{km} = \delta_{jk}\end{aligned}$$

(We used the "transposed" orthonormality condition $\left[\mathbf{R}_{-\vec{\theta}}\right]_{jm}\left[\mathbf{R}_{-\vec{\theta}}\right]_{km} = \delta_{jk}$.)
So, we see that the identity matrix has representation δ_{jk} in any frame and that the representations in different frames are related by the appropriate transformation relations.

Review of Cartesian Tensors in Classical Mechanics (cont.)

- We can demonstrate that the ϵ_{abc} Levi-Civita symbol is an isotropic rank 3 Cartesian tensor. Let's calculate the effect of the transformation rule on it:

$$\underline{(\epsilon)}'_{abc} = \left[\mathbf{R}_{-\vec{\theta}}\right]_{aj} \left[\mathbf{R}_{-\vec{\theta}}\right]_{bk} \left[\mathbf{R}_{-\vec{\theta}}\right]_{cm} \underline{(\epsilon)}_{jkm}$$

We may evaluate the above by recognizing that the "transposed" orthonormality condition on $\mathbf{R}_{-\vec{\theta}}$, $\left[\mathbf{R}_{-\vec{\theta}}\right]_{jm} \left[\mathbf{R}_{-\vec{\theta}}\right]_{km} = \delta_{jk}$, implies that the rows of $\mathbf{R}$ look like N mutually orthonormal vectors in N-dimensional space. (Here we use the term vector more loosely — we have no need to prove that these rows behave like vectors in rotated frames, we only need the fact that their component representations in a given frame looks like that of N orthonormal vectors.) Denote these "vectors" by $\vec{\mathbf{R}}^r_j$, where $\left(\vec{\mathbf{R}}^r_j\right)_k = \left[\mathbf{R}_{-\vec{\theta}}\right]_{jk}$. (The ^{r} superscript indicates we are treating the rows, rather than the columns, of $\mathbf{R}_{-\vec{\theta}}$ as vectors.) With this notation, the above product looks like $\vec{\mathbf{R}}^r_a \cdot \left(\vec{\mathbf{R}}^r_b \times \vec{\mathbf{R}}^r_c\right)$. In $N = 3$ dimensions, the expression will only be nonvanishing when the triplet abc is a cyclic or anticyclic combination; and the expression will have magnitude 1 and take the sign of the permutation (cyclic or anticyclic). These are exactly the properties of ϵ_{abc}, so we have

$$\underline{(\epsilon)}'_{abc} = \underline{(\epsilon)}_{abc}$$

So the Levi-Civita symbol is an isotropic rank 3 Cartesian tensor for $N = 3$ (and for arbitrary N, though we will not prove it here).

Note that this implies some properties of $\vec{\mathbf{M}}$:

1. When treated as a single rank 3 Cartesian tensor $\mathcal{M}$ with coordinate representation components $\underline{(\mathcal{M})}_{abc} = \left(\vec{\mathbf{M}}_a\right)_{bc} = -\underline{(\epsilon)}_{abc}$, $\mathcal{M}$ is clearly an isotropic rank 3 Cartesian tensor. For this particularly interesting case, we will take the symbol $\vec{\mathbf{M}}$ to stand for the rank 3 Cartesian tensor $\mathcal{M}$. Since $\vec{\mathbf{M}}$ is isotropic, there is no distinction between $\underline{\vec{\mathbf{M}}}$ and $\vec{\mathbf{M}}$.
2. Given a vector $\vec{\theta}$, we define the quantity $\vec{\theta} \cdot \vec{\mathbf{M}}$ by its coordinate representation in any given frame

$$\left[\vec{\theta} \cdot \vec{\mathbf{M}}\right]_{jk} = \underline{(\vec{\theta})}_a \underline{(\mathcal{M})}_{ajk}$$

   Thus, $\vec{\theta} \cdot \vec{\mathbf{M}}$ is a contraction over one index of a rank 1 and a rank 3 Cartesian tensor, yielding a rank 2 Cartesian tensor.
3. $\vec{\theta} \cdot \vec{\mathbf{M}}$ has in frames F and F' coordinate representations $\underline{\vec{\theta}} \cdot \underline{\vec{\mathbf{M}}} = \underline{\vec{\theta}} \cdot \vec{\mathbf{M}}$ and $\underline{\left(\vec{\theta} \cdot \vec{\mathbf{M}}\right)'} = \underline{\vec{\theta}}' \cdot \underline{\vec{\mathbf{M}}}' = \underline{\vec{\theta}}' \cdot \vec{\mathbf{M}}$, where the last step in each case is possible because $\vec{\mathbf{M}}$ is isotropic. Thus, only the coordinate representation of the vector $\vec{\theta}$ need be changed to write $\vec{\theta} \cdot \vec{\mathbf{M}}$ in different frames.

Active vs. Passive Transformations for Classical Cartesian Tensors

Our definition of Cartesian tensors was in terms of their transformation rules under a passive rotation — the rules tell us how to obtain, from its coordinate representation in one coordinate system F, the coordinate representation of the tensor in a new coordinate system F' that has been obtained by a rotation of the original coordinate system F by an angle $\vec{\theta}$. The transformation rule for a Cartesian tensor of rank n was Equation 15.2

$$\underline{(\mathcal{T})}'_{j_1\cdots j_n} = \left[\mathbf{R}_{-\vec{\theta}}\right]_{j_1 k_1} \cdots \left[\mathbf{R}_{-\vec{\theta}}\right]_{j_n k_n} \underline{(\mathcal{T})}_{k_1\cdots k_n}$$

(Einstein summation convention) where we have used the underline to indicate coordinate representation. The underline and the prime are outside the parentheses so it is clear we are discussing the same tensor $\mathcal{T}$; the underline indicates "coordinate representation" and the prime or lack thereof tells us which frame the coordinate representation is for. For the sake of brevity, we did not use this cumbersome but more explicit notation before, but it is necessary now. It is similar to the notation we used in discussing the Cartesian unit vectors.

We can alternatively define tensors based on active transformations, in which the coordinate system is left fixed but the tensor itself is rotated. We do this because it will be more intuitive when we go over to QM. But active transformations are more difficult to make intuitive sense of classically; it helps to think of the tensor as being attached to some physical object and for the rotation to be a rotation of the physical object. An example would be the inertia tensor of a rigid body, which rotates with the rigid body when the rigid body is rotated relative to the coordinate axes. There has been no new coordinate system created. This rotation yields a *new tensor* $\mathcal{T}'$. How do we obtain the coordinate representation of the new tensor $\mathcal{T}'$ from the coordinate representation of the original tensor $\mathcal{T}$, both in the single coordinate system F that we have referenced so far? Let's first quote the result, which is as you would expect from our discussion of passive vs. active coordinate transformations in QM:

$$\underline{(\mathcal{T}')}_{j_1\cdots j_n} = \left[\mathbf{R}_{\vec{\theta}}\right]_{j_1 k_1} \cdots \left[\mathbf{R}_{\vec{\theta}}\right]_{j_n k_n} \underline{(\mathcal{T})}_{k_1\cdots k_n} \tag{15.3}$$

Here, $\underline{(\mathcal{T}')}_{j_1\cdots j_n}$ indicates the coordinate representation of the new tensor $\mathcal{T}'$ in the frame F in which we already have the definition of the original tensor $\mathcal{T}$.

Let's prove the above, which is very similar to the earlier proof $\underline{\left(\vec{e}_j'\right)} = \mathbf{R}_{\vec{\theta}}\underline{\left(\vec{e}_j'\right)}'$. We introduce a new coordinate system F' that is rotated by the same angle $\vec{\theta}$ relative to F as the angle by which we want to rotate the tensor $\mathcal{T}$ to get the new tensor $\mathcal{T}'$. Since F' and $\mathcal{T}'$ are both obtained by the same rotation, we expect that

$$\underline{\left(\mathcal{T}'\right)}' = \underline{\left(\mathcal{T}\right)} \qquad \Longleftrightarrow \qquad \underline{\left(\mathcal{T}'\right)}'_{j_1\cdots j_n} = \underline{\left(\mathcal{T}\right)}_{j_1\cdots j_n}$$

This is the classical equivalent of unitarity, which is called orthogonality: the coordinate representation of the new tensor $\mathcal{T}'$ in the new coordinate system F' is the same as the coordinate representation of the old tensor $\mathcal{T}$ in the original coordinate system F. To some extent, this is a definition.

Given the above, we have the following based on our *passive* transformation rule:

$$\underline{\left(\mathcal{T}'\right)}'_{j_1\cdots j_n} = \left[\mathbf{R}_{-\vec{\theta}}\right]_{j_1 k_1} \cdots \left[\mathbf{R}_{-\vec{\theta}}\right]_{j_n k_n} \underline{\left(\mathcal{T}'\right)}_{k_1\cdots k_n}$$

Note that we have coordinate representations of the new tensor $\mathcal{T}'$ on both sides of the equation.

We may use orthogonality now, which lets us replace $\underline{(\mathcal{T}')}'$ with $\underline{(\mathcal{T})}$, so

$$\underline{(\mathcal{T})}_{j_1\cdots j_n} = \left[\mathbf{R}_{-\vec{\theta}}\right]_{j_1 k_1} \cdots \left[\mathbf{R}_{-\vec{\theta}}\right]_{j_n k_n} \underline{(\mathcal{T}')}_{k_1\cdots k_n}$$

Finally, we multiply both sides by $\left[\mathbf{R}_{\vec{\theta}}\right]_{m_1 j_1} \cdots \left[\mathbf{R}_{\vec{\theta}}\right]_{m_n j_n}$ and sum over the j indices. Recall that $\mathbf{R}_{-\vec{\theta}} = \mathbf{R}_{\vec{\theta}}^{-1}$, so we have

$$\begin{aligned}
\left[\mathbf{R}_{\vec{\theta}}\right]_{m_1 j_1} \cdots \left[\mathbf{R}_{\vec{\theta}}\right]_{m_n j_n} \underline{(\mathcal{T})}_{j_1\cdots j_n} &= \left[\mathbf{R}_{\vec{\theta}}\right]_{m_1 j_1} \left[\mathbf{R}_{-\vec{\theta}}\right]_{j_1 k_1} \cdots \left[\mathbf{R}_{\vec{\theta}}\right]_{m_n j_n} \left[\mathbf{R}_{-\vec{\theta}}\right]_{j_n k_n} \underline{(\mathcal{T}')}_{k_1\cdots k_n} \\
&= \delta_{m_1 k_1} \cdots \delta_{m_n k_n} \underline{(\mathcal{T}')}_{k_1\cdots k_n} \\
&= \underline{(\mathcal{T}')}_{m_1\cdots m_n} \\
\underline{(\mathcal{T}')}_{j_1\cdots j_n} &= \left[\mathbf{R}_{\vec{\theta}}\right]_{j_1 k_1} \cdots \left[\mathbf{R}_{\vec{\theta}}\right]_{j_n k_n} \underline{(\mathcal{T})}_{k_1\cdots k_n}
\end{aligned}$$

which is the expected *active* transformation rule.

Lecture 51:
Review of Classical Cartesian Tensors Continued
Eigenvectors of Classical Rotations

Date Revised: 2009/03/10
Date Given: 2009/02/27, 2009/03/02, 2009/03/04

Cartesian Tensors of Rank n as a Direct Product Hilbert Space

It should be fairly obvious that Cartesian tensors of a rank n form a Hilbert space with the real numbers as the field. The necessary "vector addition," "scalar addition and multiplication," and "scalar-vector multiplication" rules follow just as they do for rank 1 Cartesian tensors. The inner product is simply contraction over all indices, which is a scalar belonging to the real numbers based on the previous discussion:

$$\langle \mathcal{T} \, | \mathcal{S} \, \rangle = \underline{(\mathcal{T})}_{j_1 \cdots j_n} \underline{(\mathcal{S})}_{j_1 \cdots j_n}$$

Linearity and antilinearity of the inner product are proven in much the same way as for rank 1 Cartesian tensors. We will call the space of rank n Cartesian tensors $\tau^{(n)}$.

Now, what is a good basis for $\tau^{(n)}$? One reasonable basis is the set of all rank n tensors $\mathcal{E}$ whose coordinate representation in the frame F has exactly one element being 1 and all others being 0; we could label these as $\{\mathcal{E}_{j_1 \cdots j_n}\}$, $j_k = 1, \ldots, N$, $k = 1, \ldots, n$, where the indices indicate which entry is nonzero in the coordinate representation in F. That is,

$$\underline{\left(\mathcal{E}_{j_1 \cdots j_n}\right)}_{k_1 \cdots k_n} = \delta_{j_1 k_1} \cdots \delta_{j_n k_n} \tag{15.4}$$

Just as with unit vectors, a different coordinate system F' will have its own unit tensors, whose coordinate representation is simple in the F' frame:

$$\underline{\left(\mathcal{E}'_{j_1 \cdots j_n}\right)'}_{k_1 \cdots k_n} = \delta_{j_1 k_1} \cdots \delta_{j_n k_n}$$

Note the primes both inside and outside the parentheses: the former indicate that these are the unit tensors of the F' frame, and the latter indicates that we are calculating their coordinate representation in the F' frame.

Review of Cartesian Tensors in Classical Mechanics (cont.)

Let us translate tensors into our standard Hilbert space notation using kets. We will denote a tensor by $|\mathcal{T}\rangle$. Its component representation in terms of the unit tensors of a particular coordinate system is given by inner product with the appropriate unit tensor, which we know does the right thing given the definitions above:

$$\begin{aligned}\langle \mathcal{E}_{j_1\cdots j_n} | \mathcal{T} \rangle &= \underline{(\mathcal{E}_{j_1\cdots j_n})}_{k_1\cdots k_n} \underline{(\mathcal{T})}_{k_1\cdots k_n} \\ &= \delta_{j_1 k_1} \cdots \delta_{j_n k_n} \underline{(\mathcal{T})}_{k_1\cdots k_n} \\ &= \underline{(\mathcal{T})}_{j_1\cdots j_n}\end{aligned}$$

The coordinate representation in a different coordinate system F' would be given by inner product with the unit tensors of that coordinate system:

$$\begin{aligned}\langle \mathcal{E}'_{j_1\cdots j_n} | \mathcal{T} \rangle &= \underline{\left(\mathcal{E}'_{j_1\cdots j_n}\right)'}_{k_1\cdots k_n} \underline{(\mathcal{T})'}_{k_1\cdots k_n} \\ &= \delta_{j_1 k_1} \cdots \delta_{j_n k_n} \underline{(\mathcal{T})'}_{k_1\cdots k_n} \\ &= \underline{(\mathcal{T})'}_{j_1\cdots j_n}\end{aligned}$$

Review of Cartesian Tensors in Classical Mechanics (cont.)

The above basis for $\tau^{(n)}$ suggests a more fundamental way of writing $\tau^{(n)}$ and its basis. The coordinate representation of each basis element $\mathcal{E}_{j_1 \cdots j_n}$ can be written as a product of the coordinate representations of the basis elements (unit vectors) of the space of Cartesian vectors:

$$\left[\mathcal{E}_{j_1 \cdots j_n}\right]_{k_1 \cdots k_n} = \left[\vec{e}_{j_1}\right]_{k_1} \cdots \left[\vec{e}_{j_n}\right]_{k_n}$$

Since any rank n Cartesian tensor can be expanded in terms of the basis elements on the left, and any member of the space of direct products of n Cartesian vectors can be expanded in terms of the direct products of unit vectors on the right, we have a one-to-one correspondence between the basis elements of the two spaces:

$$\mathcal{E}_{j_1 \cdots j_n} \Longleftrightarrow \vec{e}_{j_1} \otimes \cdots \otimes \vec{e}_{j_n} = \bigotimes_{k=1}^{n} \vec{e}_{j_k}$$

where $\bigotimes_{k=1}^{n}$ stands for "multiple direct product", or, in ket notation

$$|\mathcal{E}_{j_1 \cdots j_n}\rangle \longleftrightarrow |\vec{e}_{j_1}\rangle \otimes \cdots \otimes |\vec{e}_{j_n}\rangle = \bigotimes_{k=1}^{n} |\vec{e}_{j_k}\rangle$$

One can check in addition that the inner product rule matches up between the two spaces, and thus the Hilbert space of tensors of rank n is equal to the direct product of the Hilbert spaces of vectors:

$$\tau^{(n)} = \tau^{(1)}_{(1)} \otimes \cdots \otimes \tau^{(1)}_{(n)} = \bigotimes_{k=1}^{n} \tau^{(1)}_{(k)}$$

where there are n elements in the direct product and where the subscripts refer to which of the n elements is being referenced. The idea of using direct products of unit vectors in $\tau^{(1)}$ as the basis for $\tau^{(n)}$ is thus the natural thing to do.

Review of Cartesian Tensors in Classical Mechanics (cont.)

With the abstract idea of a Hilbert space, we also have the concept of operators that are "more than" their matrix representations in a specific basis. In particular, we must abstract the rotation operator from the rotation matrix we have been working with. Of course, operators are usually defined in terms of their matrix representations in a particular basis. So, we simply **define** the rotation operator $R^{(n)}(\vec{\theta})$ acting on the Hilbert space $\tau^{(n)}$ to be the operator whose matrix representation in the basis of the unit tensors of a particular coordinate frame F is the necessary rotation matrices that act on the coordinate representation of tensors in that frame. That is, $R^{(n)}(\vec{\theta})$ is the operator with matrix representation given by

$$\langle \mathcal{E}_{j_1 \cdots j_n} | R^{(n)}(\vec{\theta}) | \mathcal{E}_{k_1 \cdots k_n} \rangle = \left[\mathbf{R}_{\vec{\theta}}\right]_{j_1 k_1} \cdots \left[\mathbf{R}_{\vec{\theta}}\right]_{j_n k_n} \tag{15.5}$$

This form does exactly what one wants to the coordinate representations of tensors. Suppose we want to do an active rotation on a tensor $\mathcal{T}$ with coordinate representation $\{\underline{(\mathcal{T})}_{j_1\cdots j_n}\}$ in F to get a new tensor $\mathcal{T}'$ with coordinate representation $\{\underline{(\mathcal{T}')}_{j_1\cdots j_n}\}$ in F. Let's do this using the above operator, and project onto the basis of unit tensors of F to recover the coordinate representation of $\mathcal{T}'$:

$$\begin{aligned}
\underline{(\mathcal{T}')}_{j_1\cdots j_n} &= \langle \mathcal{E}_{j_1\cdots j_n} | \mathcal{T}' \rangle \\
&= \langle \mathcal{E}_{j_1\cdots j_n} | R^{(n)}(\vec{\theta}) | \mathcal{T} \rangle \\
&= \langle \mathcal{E}_{j_1\cdots j_n} | R^{(n)}(\vec{\theta}) \sum_{k_1\cdots k_n} \underline{(\mathcal{T})}_{k_1\cdots k_n} | \mathcal{E}_{k_1\cdots k_n} \rangle \\
&= \sum_{k_1\cdots k_n} \underline{(\mathcal{T})}_{k_1\cdots k_n} \langle \mathcal{E}_{j_1\cdots j_n} | R^{(n)}(\vec{\theta}) | \mathcal{E}_{k_1\cdots k_n} \rangle \\
&= \sum_{k_1\cdots k_n} \left[\mathbf{R}_{\vec{\theta}}\right]_{j_1 k_1} \cdots \left[\mathbf{R}_{\vec{\theta}}\right]_{j_n k_n} \underline{(\mathcal{T})}_{k_1\cdots k_n}
\end{aligned}$$

as desired.

We can also see that this definition of $R^{(n)}(\vec{\theta})$ gives us the right passive transformation behavior. If $\{|\mathcal{E}_{j_1\cdots j_n}\rangle\}$ are the unit tensors of a frame F and $\{|\mathcal{E}'_{j_1\cdots j_n}\rangle\}$ are the unit tensors of a different frame F', with F' obtained by a rotation of F by $\vec{\theta}$, then we can obtain the coordinate representation of a tensor $\mathcal{T}$ in the frame F', $\{\underline{(\mathcal{T})}'_{j_1\cdots j_n}\}$ from its representation in the frame F, $\{\underline{(\mathcal{T})}_{j_1\cdots j_n}\}$ using the operator $R^{(n)}(\vec{\theta})$. First note that the definition of F' relative to F implies that the unit tensors of F' are obtained by active rotation of the unit tensors of F:

$$|\mathcal{E}'_{j_1\cdots j_n}\rangle = R^{(n)}(\vec{\theta})|\mathcal{E}_{j_1\cdots j_n}\rangle$$

Then,
$$\begin{aligned}
\underline{(\mathcal{T}')}_{j_1\cdots j_n} &= \langle\mathcal{E}'_{j_1\cdots j_n}|\mathcal{T}\rangle \\
&= \langle\mathcal{E}_{j_1\cdots j_n}|\left[R^{(n)}(\vec{\theta})\right]^{\dagger}|\mathcal{T}\rangle \\
&= \sum_{k_1\cdots k_n}\underline{(\mathcal{T})}_{k_1\cdots k_n}\langle\mathcal{E}_{j_1\cdots j_n}|R^{(n)}(-\vec{\theta})|\mathcal{E}_{k_1\cdots k_n}\rangle \\
&= \sum_{k_1\cdots k_n}\left[\mathbf{R}_{-\vec{\theta}}\right]_{j_1k_1}\cdots\left[\mathbf{R}_{-\vec{\theta}}\right]_{j_nk_n}\underline{(\mathcal{T})}_{k_1\cdots k_n}
\end{aligned}$$

as expected.

Our definition of $R^{(n)}(\vec{\theta})$ has the nice feature of yielding a sensible direct product breakdown:

$$\begin{aligned}
\langle \mathcal{E}_{j_1\cdots j_n} | R^{(n)}(\vec{\theta}) | \mathcal{E}_{k_1\cdots k_n} \rangle &= \left[\mathbf{R}_{\vec{\theta}}\right]_{j_1 k_1} \cdots \left[\mathbf{R}_{\vec{\theta}}\right]_{j_n k_n} \\
&= \langle \vec{e}_{j_1} | R^{(1)}(\vec{\theta}) | \vec{e}_{k_1} \rangle \cdots \langle \vec{e}_{j_n} | R^{(1)}(\vec{\theta}) | \vec{e}_{k_n} \rangle \\
&= \left(\langle \vec{e}_{j_1} | \otimes \cdots \otimes \langle \vec{e}_{j_n} | \right) \left[R^{(1)}_{(1)}(\vec{\theta}) \otimes \cdots \otimes R^{(1)}_{(n)}(\vec{\theta}) \right] \left(| \vec{e}_{k_1} \rangle \otimes \cdots \otimes | \vec{e}_{k_n} \rangle \right) \\
&= \langle \mathcal{E}_{j_1\cdots j_n} | \left[R^{(1)}_{(1)}(\vec{\theta}) \otimes \cdots \otimes R^{(1)}_{(n)}(\vec{\theta}) \right] | \mathcal{E}_{k_1\cdots k_n} \rangle
\end{aligned}$$

$$\iff \qquad R^{(n)}(\vec{\theta}) = R^{(1)}_{(1)}(\vec{\theta}) \otimes \cdots \otimes R^{(1)}_{(n)}(\vec{\theta}) = \bigotimes_{k=1}^{n} R^{(1)}_{(k)}(\vec{\theta}) \tag{15.6}$$

Along with the generalized rotation operator, we should define a generalized generator $i\,\hbar\,\vec{M}^{(n)}$. The above direct product breakdown shows us how to define it. Let's first do the $n = 1$ case by defining

$$\langle \vec{e}_j \,|\, i\,\hbar\,\vec{M}^{(1)} | \vec{e}_k \,\rangle \equiv \Big[i\,\hbar\,\vec{\mathbf{M}} \Big]_{jk} \qquad \Longleftrightarrow \qquad \langle \vec{e}_j \,|\, i\,\hbar\,M_a^{(1)} | \vec{e}_k \,\rangle \equiv [\, i\,\hbar\,\mathbf{M}_a \,]_{jk} \tag{15.7}$$

which implies

$$\begin{aligned} \langle \vec{e}_j \,|\, \exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot i\,\hbar\,\vec{M}^{(1)} \right) | \vec{e}_k \,\rangle &= \left[\exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot i\,\hbar\,\vec{\mathbf{M}} \right) \right]_{jk} \\ &= \left[\mathbf{R}_{\vec{\theta}} \right]_{jk} = \langle \vec{e}_j \,|\, R^{(1)}(\vec{\theta}) | \vec{e}_k \,\rangle \end{aligned}$$

In going from the first expression to the second expression in the above, we skipped over the steps where we wrote out the power series expansion of the exponential, inserted completeness in between each power of $\vec{M}^{(1)}$ to replace $\vec{M}^{(1)}$ with matrix elements of $\vec{\mathbf{M}}$, and then recollapsed the power series to be apower series in the $\vec{\mathbf{M}}$ matrix rather than the $\vec{M}^{(1)}$ operator. Thus, we now have an operator version of the generator for $n = 1$.

Now, let's define the generalized generator for $n > 1$ in terms of $\vec{M}^{(1)}$:

$$
\begin{aligned}
i\,\hbar\,\vec{M}^{(n)} &\equiv i\,\hbar \sum_{k=1}^{n} \vec{M}^{(1)}_{(k)} \otimes \bigotimes_{p\neq k}^{n} I_{(p)} \qquad (15.8)\\
&= i\,\hbar\Big[\vec{M}^{(1)}_{(1)} \otimes I_{(2)} \otimes \cdots \otimes I_{(n)} + I_{(1)} \otimes \vec{M}^{(1)}_{(2)} \otimes I_{(3)} \otimes \cdots \otimes I_{(n)}\\
&\qquad + I_{(1)} \otimes \cdots \otimes I_{(n-1)} \otimes \vec{M}^{(1)}_{(n)}\Big]\\
i\,\hbar\,M^{(n)}_{a} &\equiv i\,\hbar \sum_{k=1}^{n} M^{(1)}_{(k),a} \otimes \bigotimes_{p\neq k}^{n} I_{(p)}
\end{aligned}
$$

where the $_{(k)}$ and $_{(p)}$ indicates which of the n factor spaces of Cartesian vectors the operators act in.

We can see that the above definition is consistent with our definition of the abstract rotation operator:

$$\begin{aligned}
\exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot i\,\hbar\,\vec{M}^{(n)}\right) &= \exp\left(-\frac{i}{\hbar}\sum_{k=1}^{n}\vec{\theta}\cdot i\,\hbar\,\vec{M}^{(1)}_{(k)}\otimes\bigotimes_{p\neq k}^{n} I_{(p)}\right) \\
&= \prod_{k=1}^{n}\left[\exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot i\,\hbar\,\vec{M}^{(1)}_{(k)}\right)\otimes\bigotimes_{p\neq k}^{n} I_{(p)}\right] \\
&= \bigotimes_{k=1}^{n}\exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot i\,\hbar\,\vec{M}^{(1)}_{(k)}\right) = \bigotimes_{k=1}^{n} R^{(1)}_{(k)}(\vec{\theta}) = R^{(n)}(\vec{\theta})
\end{aligned}$$

as desired, so our generalized generator definition is correct. Note that, in going from the second line to the third line, we made use of the fact that each term in the sum in the argument of the exponential commutes with every other term in the sum because each term has nontrivial action in only one of the n factor spaces and each one acts in a different factor space.

Review of Cartesian Tensors in Classical Mechanics (cont.)

There is one aspect of the above that we need to explain. It appears that our definition of the matrix elements of $R^{(n)}(\vec{\theta})$ is the same regardless of the frame for which the $\{\mathcal{E}_{j_1 \ldots j_n}\}$ are the unit tensors. How can the representation of an operator be independent of the basis chosen?

The answer is that the representation is not independent of the frame because the $\mathbf{R}_{\vec{\theta}}$ that provide the coordinate representation for $R^{(n)}(\vec{\theta})$ depend on the coordinate representation of $\vec{\theta}$, which depends on one's choice of coordinate frame F. The formulae we have given that relate $R^{(n)}(\vec{\theta})$ and $\mathbf{R}_{\vec{\theta}}$ are completely correct and general, but they have a coordinate frame dependence through the coordinate frame dependence of the representation of $\vec{\theta}$.

A similar issue occurs for the generator $\vec{M}^{(n)}$, too. The operator $\vec{M}^{(n)}$ is not frame-independent because it assumes a set of unit vectors that define the three directions for component operators in $\vec{M}^{(n)}$. The fact that the formulae that define $\vec{M}^{(n)}$ are not dependent on the frame just implies that one always calculates the version of $\vec{M}^{(n)}$ that is tied to the coordinate frame one is working in. Here, we make a distinction between frame independence and basis independence. A coordinate frame defines a set of Cartesian unit vectors $\{\vec{e}_j\}$. But we need not use those as a basis for the Hilbert space, though we may still work in that coordinate frame. This point will become clearer below.

Choosing a More Convenient Basis for $\tau^{(1)}$: The Eigenvector-Eigenvalue Problem of Rotations for Rank 1 Cartesian Tensors

We now know that $\tau^{(n)}$ is a Hilbert space, and we have a basis that is composed of direct products of the natural basis elements of $\tau^{(1)}$, the Cartesian unit vectors. That's fine, but we are bringing the idea of tensors up because we are interested in setting up quantum mechanical states with reasonable properties under coordinate system rotations. We know from our study of orbital angular momentum that Cartesian coordinates are not really the natural way to discuss rotations. We found in Section 14 that the structure of the Hilbert space of QM states of a particle in three spatial dimensions breaks down quite cleanly if we consider the properties of the states under rotations by finding the eigenvectors and eigenvalues of the L^2 and L_z operators. We found a basis of eigenvectors and eigenvalues of L^2 and L_z, the $\{|\ell, m\rangle\}$, and we saw that the space breaks down into a direct sum of degenerate subspaces of L^2:

$$\mathbb{V} = \mathbb{V}^{(0)} \oplus \mathbb{V}^{(1)} \oplus \mathbb{V}^{(2)} \oplus \cdots$$

(When we considered the more general J^2 and J_z operators, we also found half-integer j values, but we don't need to consider the most general case right now; this discussion is mainly for motivation.). Each of these subspaces is invariant under rotations. Moreover, the $|\ell, m\rangle$ states are themselves invariant under rotations about z because such rotations are generated by L_z because they are eigenvectors of L_z. That makes them eigenvectors of $T(\vec{\theta}) = \exp\left(-\frac{i}{\hbar}\,\theta_z L_z\right)$ with eigenvalues $e^{-i\,m\,\theta_z}$, $m = -\ell,\ -\ell+1,\ \ldots,\ \ell-1,\ \ell$.

We did not emphasize the latter point much, but it provides important motivation here. We know that $\vec{M}^{(1)}$ generates classical rotations (recall, $R^{(1)}(\vec{\theta}) = \exp(\vec{\theta} \cdot \vec{M}^{(1)})$), that the components of $\vec{M}^{(1)}$ satisfy a commutation relation similar to the one the components of $\vec{L}$ satisfy, and that $[M^2]^{(1)}$ commutes with each component of $\vec{M}^{(1)}$. (We are using the matrix-representation-free operators $\vec{M}^{(1)}$ and $[M^2]^{(1)}$ here because we want to avoid getting tied to a basis whenever possible.) So, we are led to ask the question: is there a basis for $\tau^{(1)}$ that consists of eigenvectors of $[M^2]^{(1)}$ and $M_z^{(1)}$? These eigenvectors would presumably simplify the behavior of $\tau^{(1)}$ under rotations in the same way that the $|\ell, m\rangle$ basis simplified the behavior of the the Hilbert space of states of a particle in three spatial dimensions under rotations.

Note that, just because we are not yet picking a basis, we have indeed picked a coordinate frame because we must choose Cartesian unit vectors before the operator $\vec{M}^{(1)}$ makes sense.

Before finding these eigenvectors explicitly, we can predict the results. The commutation relations of the components of $\vec{M}^{(1)}$ imply that

$$[i\,\hbar\,M_a^{(1)}, i\,\hbar\,M_b^{(1)}] = \epsilon_{abc}\,i\,\hbar\,M_c^{(1)} \qquad [-\hbar^2[M^2]^{(1)}, i\,\hbar\,M_a^{(1)}] = 0$$

The components of $i\,\hbar\,\vec{M}^{(1)}$ satisfy the same generic commutation relations as the components of $\vec{J}$ and $\vec{L}$. Thus, our operator method analysis of the eigenvectors and eigenvalues of J^2 and J_z applies here: we are assured that there is a basis of eigenvectors of $-\hbar^2[M^2]^{(1)}$ and $i\,\hbar\,M_z^{(1)}$ that we can label $|\vec{e}_m^{(j)}\rangle$ and that satisfy

$$-\hbar^2[M^2]^{(1)}|\vec{e}_m^{(j)}\rangle = \hbar^2 j\,(j+1)\,|\vec{e}_m^{(j)}\rangle \qquad i\,\hbar\,M_z^{(1)}|\vec{e}_m^{(j)}\rangle = \hbar\,m\,|\vec{e}_m^{(j)}\rangle$$

$$j = \frac{k}{2} \quad k \text{ any integer} \quad m = -j, -(j-1), \cdots, j-1, j$$

However, that discussion only said the above values were allowed; it did not say they had to exist. For example, in the Hilbert space of states of a particle in three spatial dimensions, we have shown that the particulars of the problem (*i.e.*, the specific representation of the angular momentum operators, which yield the specific eigenvalue-eigenvector differential equations) imply that only the integral values of j exist. What $\vec{e}_m^{(j)}$ states exist in $\tau^{(1)}$?

To find what the allowed values of j are, we do the obvious thing, which is to write down the eigenvector-eigenvalue equations of $-\hbar^2[M^2]^{(1)}$:

$$-\hbar^2[M^2]^{(1)}|\vec{e}_m^{(j)}\rangle = \hbar^2 j\,(j+1)\,|\vec{e}_m^{(j)}\rangle$$

We know $\mathbf{M}^2 = -2\,\mathbf{I}$, so we may infer the basis-independent statement $[M^2]^{(1)} = -2I$. This gives

$$2\,\hbar^2\,I|\vec{e}_m^{(j)}\rangle = \hbar^2 j\,(j+1)\,|\vec{e}_m^{(j)}\rangle$$

Thus, the $-\hbar^2[M^2]^{(1)}$ equation tells us $j = 1$. We may now write $\vec{e}_m^{(1)}$ instead of $\vec{e}_m^{(j)}$.

Next, let's solve the $i\,\hbar\,M_z^{(1)}$ equation. Because we don't have such a simple form for $i\,\hbar\,M_z^{(1)}$ as we had for $-\hbar^2[M^2]^{(1)}$, we must project this equation onto a basis, so we use the obvious basis of Cartesian unit vectors, for which we know the representation of $M_z^{(1)}$ is $\mathbf{M}_z$

$$i\,\hbar\,M_z^{(1)}|\vec{e}_m^{(1)}\rangle = m\,\hbar\,|\vec{e}_m^{(1)}\rangle$$

$$i\,\hbar\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \underline{\left(\vec{e}_m^{(1)}\right)_1} \\ \underline{\left(\vec{e}_m^{(1)}\right)_2} \\ \underline{\left(\vec{e}_m^{(1)}\right)_3} \end{bmatrix} = m\,\hbar\begin{bmatrix} \underline{\left(\vec{e}_m^{(1)}\right)_1} \\ \underline{\left(\vec{e}_m^{(1)}\right)_2} \\ \underline{\left(\vec{e}_m^{(1)}\right)_3} \end{bmatrix} \quad \text{with} \quad |\vec{e}_m^{(1)}\rangle \underset{\vec{e}}{\longleftrightarrow} \begin{bmatrix} \underline{\left(\vec{e}_m^{(1)}\right)_1} \\ \underline{\left(\vec{e}_m^{(1)}\right)_2} \\ \underline{\left(\vec{e}_m^{(1)}\right)_3} \end{bmatrix}$$

where we write $\underset{\vec{e}}{\longleftrightarrow}$ to indicate the coordinate representation in the standard Cartesian coordinate system, or, equivalently, matrix representation in the basis of Cartesian unit vectors of $\tau^{(1)}$ for the coordinate frame we are working in.

Review of Cartesian Tensors in Classical Mechanics (cont.)

Finding the roots of the characteristic polynomial for the $i\hbar M_z^{(1)}$ eigenvalue problem shows us that the expected values m, $m = -1, 0, 1$, are indeed realized:

$$m\hbar\left(m^2\hbar^2 - \hbar^2\right) = 0 \qquad \Longrightarrow \qquad m = 0, \pm 1$$

The eigenvectors are

$$\vec{e}_1^{(1)} \underset{\vec{e}}{\longleftrightarrow} \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \\ 0 \end{bmatrix} \qquad \vec{e}_0^{(1)} \underset{\vec{e}}{\longleftrightarrow} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad \vec{e}_{-1}^{(1)} \underset{\vec{e}}{\longleftrightarrow} \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \\ 0 \end{bmatrix}$$

Because two of the eigenvectors are complex, we need to expand $\tau^{(1)}$ to allow complex coefficients. There are three reasons to just go ahead and do this rather than worrying unduly about the fact that classical vectors are not complex. First, though we have a basis of complex vectors, we can write any real vector in terms of them (because they are a basis for the complex vector space in which the real vectors reside), and a coordinate system rotation must leave a real vector as a real vector. Second, the definition of a vector by its behavior under spatial rotations does not require that the components be real; that is, there is nothing fundamental in the nature of a vector about having real coefficients except that the physical vectors we are familiar with from classical mechanics satisfy this requirement. Third, when we take this over to quantum mechanics, we will naturally need complex coefficients.

Since the above vectors are eigenvectors of $i\,\hbar\,M_z^{(1)}$, we thus know that they are eigenvectors of the operator that rotates three-dimensional vectors about the z axis,

$$R^{(1)}(\theta_z\widehat{z}) = \exp\left(\theta_z M_z^{(1)}\right) = \exp\left(-\frac{i}{\hbar}\,\theta_z\left(i\,\hbar\,M_z^{(1)}\right)\right)$$

with eigenvalues $e^{-i\,m\,\theta_z}$, 1, and $e^{i\,m\,\theta_z}$. That is, these three vectors are invariant under rotations about z. We did not emphasize this at the time, but the same fact held for the $|j,m\rangle$ states when acted upon by the operator for rotations about z axis, $T(\theta_z\widehat{z})$.

We also note the obvious fact that the most general rotation operator on this space

$$R^{(1)}(\vec{\theta}) = \exp\left(\vec{\theta}\cdot\vec{M}^{(1)}\right) = \exp\left(-\frac{i}{\hbar}\,\vec{\theta}\cdot\left(i\,\hbar\,\vec{M}^{(1)}\right)\right)$$

acts on vectors in this space $\tau^{(1)}$ and produces vectors in this space $\tau^{(1)}$. This is identical to the way in which $T(\vec{\theta})$ does not change the j value of $|j,m\rangle$ states: it keeps vectors in $\mathbb{V}^{(j)}$ in $\mathbb{V}^{(j)}$, or, equivalently, $\mathbb{V}^{(j)}$ is invariant or closed under the action of $T(\vec{\theta})$.

Review of Cartesian Tensors in Classical Mechanics (cont.)

Finally, since we have a new basis for the space that $-i\,\hbar\,M_z^{(1)}$ and $-\hbar^2[M^2]^{(1)}$ work on, the basis of their eigenvectors, let's write their matrix representation in this basis:

$$i\,\hbar\,M_z^{(1)} \xleftrightarrow[\vec{e}^{(1)}]{} \hbar \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \xleftrightarrow[|j,m\rangle]{} J_z^{(1)}$$

$$-\hbar^2[M^2]^{(1)} \xleftrightarrow[\vec{e}^{(1)}]{} 2\,\hbar^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xleftrightarrow[|j,m\rangle]{} \left[J^2\right]^{(1)}$$

That is, we see that the matrix representations of $-i\,\hbar\,M_z^{(1)}$ and $-\hbar^2[M^2]^{(1)}$ in the basis of their eigenvectors $\{\vec{e}_m^{(1)}\}$ are identical to the matrix representations of the J_z and J^2 operators in the $|j,m\rangle$ basis for the $\mathbb{V}^{(1)}$ space. You can check that, if one writes matrix representations of $M_x^{(1)}$ and $M_y^{(1)}$ in this same basis, those representations would also match up to those of $J_x^{(1)}$ and $J_y^{(1)}$ and thus the correspondence $i\,\hbar\,\vec{M}^{(1)} \leftrightarrow \vec{J}^{(1)}$ is perfect. Thus, our operators $-i\,\hbar\,\vec{M}^{(1)}$ and $-\hbar^2[M^2]^{(1)}$ are really just analogues of $\vec{J}$ and J^2 for the space $\tau^{(1)}$. This is an extremely interesting statement — it says that our generalized QM formalism for rotation operations, the "operator method" that derived only from the commutation relation $[J_a, J_b] = i\,\hbar\,J_c$, also works for classical rank 1 Cartesian tensors. This suggests that it will be reasonable to define quantum mechanical states for particles with spin by making use of classical tensors in some way.

Review of Cartesian Tensors in Classical Mechanics (cont.)

Let us make use of these correspondences to simplify our notation. First of all, the above correspondence tells us

$$\tau^{(1)} = \mathbb{V}^{(1)}$$

While the two spaces may have arisen out of different physics, they are, for all intents and purposes, identical. For the sake of clarity, we will continue to make a distinction between them wherein we will refer to the space as $\tau^{(1)}$ when we use the basis of Cartesian unit vectors and as $\mathbb{V}^{(1)}$ when we use the basis of eigenvectors of $-i\,\hbar\,M_z^{(1)}$ and $-\hbar^2[M^2]^{(1)}$.

We also realize that $i\,\hbar\,\vec{M}^{(1)}$ and $\vec{J}^{(1)}$ are the same operator on this space. We defined $i\,\hbar\,\vec{M}^{(1)}$ through its matrix representation in the Cartesian unit vector $\{\vec{e}_j\}$ basis via $i\,\hbar\,\vec{M}^{(1)} \underset{\vec{e}}{\longleftrightarrow} i\,\hbar\,\vec{\mathbf{M}}$, but our correspondence above implies that the matrix representation of $i\,\hbar\,\vec{M}^{(1)}$ in the $\{\vec{e}_m^{(1)}\}$ basis is the same as that of $\vec{J}^{(1)}$ in the $\{|j,m\rangle\}$ basis. We will now use $\vec{J}^{(1)}$ in all cases.

Review of Cartesian Tensors in Classical Mechanics (cont.)

In order to clearly distinguish between operators and matrix representations, we define the matrix $\vec{\mathbf{J}}^{(1)}$ to be the above matrix representation of $i\,\hbar\,\vec{M}^{(1)}$ and $\vec{J}^{(1)}$. That is,

$$i\,\hbar\,\vec{M}^{(1)} \xleftrightarrow[\vec{e}^{(1)}]{} \vec{\mathbf{J}}^{(1)} \xleftrightarrow[|j,m\rangle]{} \vec{J}^{(1)} \qquad \text{with}$$

$$\mathbf{J}_z^{(1)} = \hbar \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \mathbf{J}_x^{(1)} = \frac{\hbar}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \mathbf{J}_y^{(1)} = \frac{i\,\hbar}{\sqrt{2}} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

Then, we may write the matrix representations of the classical vector rotation operator $R^{(1)}(\vec{\theta})$ in the conventional Cartesian $\{\vec{e}_j\}$ basis and the rotation eigenvector $\{\vec{e}_m^{(1)}\}$ basis

$$\langle \vec{e}_j \,|R^{(1)}(\vec{\theta})|\vec{e}_k\,\rangle = \left[\exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot i\,\hbar\,\vec{\mathbf{M}}\right)\right]_{jk} = \left[\mathbf{R}_{\vec{\theta}}\right]_{jk}$$

$$\langle \vec{e}_{m_1}^{(1)} \,|R^{(1)}(\vec{\theta})|\vec{e}_{m_2}^{(1)}\,\rangle = \left[\exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot\,\vec{\mathbf{J}}^{(1)}\right)\right]_{m_1 m_2} \equiv \left[\mathbf{R}_{\vec{\theta}}^{(1)}\right]_{m_1 m_2}$$

where we **define** $\mathbf{R}_{\vec{\theta}}^{(1)}$ to be the matrix representation of $R^{(1)}(\vec{\theta})$ in the $\{\vec{e}_m^{(1)}\}$ basis; it is also obtained by diagonalizing $\mathbf{R}_{\vec{\theta}}$.

We recognize from the above that the matrix representation of $R^{(1)}(\vec{\theta})$ in the $\{\vec{e}_m^{(1)}\}$ basis is the same as that of $T^{(1)}(\vec{\theta})$ in the $\{|j,m\rangle\}$ basis. Since these two bases are the same when one recognizes $\tau^{(1)} = \mathbb{V}^{(1)}$, it follows that $R^{(1)}(\vec{\theta}) = T^{(1)}(\vec{\theta})$. This also follows directly from $i\,\hbar\,\vec{M}^{(1)} = \vec{J}^{(1)}$, $R^{(1)}(\vec{\theta}) = \exp\left(-\frac{i}{\hbar}\,\vec{\theta}\cdot i\,\hbar\,\vec{M}^{(1)}\right)$, and $T^{(1)}(\vec{\theta}) = \exp\left(-\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{J}^{(1)}\right)$.

As with $i\,\hbar\,\vec{M}^{(1)}$ and $\vec{J}^{(1)}$, we will use $R^{(1)}(\vec{\theta})$ when discussing $\tau^{(1)}$ in terms of its Cartesian unit vector basis and we will use $T^{(1)}(\vec{\theta})$ when discussing it in the $\{\vec{e}_m^{(1)}\}$ basis. We will always use $\mathbf{R}^{(1)}_{\vec{\theta}}$ for the the matrix representation in the latter basis because we have no prior notation for this matrix.

Lecture 52:
Classical Spherical Tensors

Date Revised: 2009/03/04
Date Given: 2009/03/02, 2009/03/04

Spherical Tensors

The above example shows us that there is a different way of writing rank 1 Cartesian tensors that makes their properties under rotations far clearer because the new basis gives us the eigenvectors of the rotation operator. Can we generalize this idea of "simple behavior" under rotations?

The first step to generalizing is figuring out what the generalized rotation operation should be. Our vector example immediately suggests that we should look at the matrix representations of the $\vec{J}$ and J^2 operators that we have found for $j \neq 1$. We know how to construct rotation operators based on these other j matrix representations.

Why are Cartesian tensors of rank n not the right objects to consider? We will see later that it is because they are reducible, while the spherical tensors we are about to define are irreducible. It would not be incorrect to use Cartesian tensors, but it would not have the same simplicity. You cannot see this right now. For now, we work under the assumption that generalizing the rotation operator generated by $\vec{J}^{(1)}$ is the right way to go.

Spherical Tensors in Classical Mechanics (cont.)

We thus **define** a *rank j spherical tensor* to be an object $\mathcal{T}^{(j)}$ whose coordinate representation $\underline{(\mathcal{T})}^{(j)}$ in a coordinate system consists of a set of $2j+1$ numbers $\{T^{(j)}_m\}$ (or $\left\{\underline{\left(\mathcal{T}^{(j)}\right)}_m\right\}$), with $m = -j, -j+1, \ldots, j$, that satisfy the transformation rule under *passive* transformations

$$\underline{\left(\mathcal{T}^{(j)}\right)}' = \mathbf{R}^{(j)}_{-\vec{\theta}}\underline{\left(\mathcal{T}^{(j)}\right)} = \exp\left(-\frac{i}{\hbar}\left(-\vec{\theta}\right)\cdot\vec{\mathbf{J}}^{(j)}\right)\underline{\left(\mathcal{T}^{(j)}\right)} \tag{15.9}$$

$$\underline{\left(\mathcal{T}^{(j)}\right)}'_m = \sum_{q=-j}^{j}\left[\mathbf{R}^{(j)}_{-\vec{\theta}}\right]_{mq}\underline{\left(\mathcal{T}^{(j)}\right)}_q = \sum_{q=-j}^{j}\left[\exp\left(-\frac{i}{\hbar}\left(-\vec{\theta}\right)\cdot\vec{\mathbf{J}}^{(j)}\right)\right]_{mq}\underline{\left(\mathcal{T}^{(j)}\right)}_q$$

where $\underline{\left(\mathcal{T}^{(j)}\right)}$ is the coordinate representation of the spherical tensor $\mathcal{T}^{(j)}$ in the original frame F and $\underline{\left(\mathcal{T}^{(j)}\right)}'$ is the coordinate representation of the same spherical tensor in the coordinate system F' that is obtained by rotation by the angle $\vec{\theta}$ relative to F. We write $\vec{\mathbf{J}}^{(j)}$ to indicate the matrix representation of the operator $\vec{J}^{(j)}$ in the $|j,m\rangle$ basis; remember, $\vec{J}^{(j)}$ is a representation-free operator, not a matrix! Similarly, $\mathbf{R}^{(j)}_{\vec{\theta}}$ is the matrix representation in the $|j,m\rangle$ basis of the operator $T^{(j)}(\vec{\theta})$ defined in Equation 14.8. Recall that, $\mathbf{R}^{(1)}_{\vec{\theta}}$ is also the diagonalized version of our Cartesian tensor rotation matrix $\mathbf{R}_{\vec{\theta}}$. Note that, by dint of their definition in terms of the Hermitian $\vec{\mathbf{J}}^{(j)}$ matrices, the $\mathbf{R}^{(j)}_{\vec{\theta}}$ matrices are unitary.

Spherical Tensors in Classical Mechanics (cont.)

The analogous *active* transformation of spherical tensors is in direct analogy to the relation between passive and active transformations for Cartesian tensors: we just change the sign on the rotation angle and replace $\underline{\left(\mathcal{T}^{(j)}\right)'}$ with $\underline{\left(\mathcal{T}^{(j)\,\prime}\right)'}$:

$$\underline{\left(\mathcal{T}^{(j)\,\prime}\right)} = \mathbf{R}^{(j)}_{\vec{\theta}}\,\underline{\left(\mathcal{T}^{(j)}\right)} = \exp\left(-\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{\mathbf{J}}^{(j)}\right)\underline{\left(\mathcal{T}^{(j)}\right)} \tag{15.10}$$

$$\underline{\left(\mathcal{T}^{(j)\,\prime}\right)}_{m} = \sum_{q=-j}^{j}\left[\mathbf{R}^{(j)}_{\vec{\theta}}\right]_{mq}\underline{\left(\mathcal{T}^{(j)}\right)}_{q} = \sum_{q=-j}^{j}\left[\exp\left(-\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{\mathbf{J}}^{(j)}\right)\right]_{mq}\underline{\left(\mathcal{T}^{(j)}\right)}_{q}$$

We will not prove this explicitly because the proof is identical in technique to the one we made for Cartesian tensors (and, in fact, is easier because we always have only a single rotation matrix for spherical tensors).

We know from finding the eigenvectors of $i\,\hbar\,M_z^{(1)}$ and $-\hbar^2[M^2]^{(1)}$ that rank 1 Cartesian and spherical tensors are just different matrix representations of the same objects: we call these objects rank 1 Cartesian tensors when we write their coordinate representations in terms of the Cartesian unit vectors and consider the transformation rules for that coordinate representation under rotations, and we call them rank 1 spherical tensors when we write their coordinate representations in terms of the eigenvectors of $-i\,\hbar\,M_z^{(1)}$ and $-\hbar^2[M^2]^{(1)}$ and consider the transformation rules for that coordinate representation under transformations.

Spherical Tensors in Classical Mechanics (cont.)

Our rank j spherical tensors are, by definition, the space $\mathbb{V}^{(j)}$ because we have defined them by requiring their transformation rules under rotations are the same as that of the elements of the $\mathbb{V}^{(j)}$ space. We may define unit spherical tensors $\{\vec{e}_m^{(j)}\}$ of a coordinate frame F to be the spherical tensors whose coordinate representation in that frame are

$$\underline{\left(\vec{e}_m^{(j)}\right)}_n = \delta_{mn}$$

Because the space of spherical tensors is $\mathbb{V}^{(j)}$, there is a one-to-one relationship between these spherical unit tensors and the $|j,m\rangle$ basis for this j:

$$\vec{e}_m^{(j)} = |j,m\rangle$$

Note the use of an equals sign rather than a representation correspondence — spherical unit tensors are the $|j,m\rangle$ basis elements and vice versa. There is no reason to define Hilbert space basis kets $|\vec{e}_m^{(j)}\rangle$ as we did for Cartesian tensors because we already have them in the form of the $|j,m\rangle$.

Spherical Tensors in Classical Mechanics (cont.)

The above implies that

$$\mathcal{T}^{(j)} = \sum_{m=-j}^{j} \underline{(\mathcal{T}^{(j)})}_m \vec{e}_m^{(j)} \quad \Longleftrightarrow \quad |\mathcal{T}^{(j)}\rangle = \sum_{m=-j}^{j} \underline{(\mathcal{T}^{(j)})}_m |j,m\rangle$$

We can see that the basis-free way of writing a rotation operation is exactly as we did for $|j,m\rangle$ states:

$$|\mathcal{T}^{(j)\,\prime}\rangle = T^{(j)}(\vec{\theta})|\mathcal{T}^{(j)}\rangle$$

by simply projecting the above onto the $|j,m\rangle$ basis:

$$\langle j,m|\mathcal{T}^{(j)\,\prime}\rangle = \langle j,m|T^{(j)}(\vec{\theta}) \sum_{q=-j}^{j} |j,q\rangle\langle j,q|\mathcal{T}^{(j)}\rangle$$

$$\underline{\left(\mathcal{T}^{(j)\,\prime}\right)}_m = \sum_{q=-j}^{j} \left[\mathbf{R}_{\vec{\theta}}^{(j)}\right]_{mq} \underline{\left(\mathcal{T}^{(j)}\right)}_q$$

Spherical Tensors in Classical Mechanics (cont.)

Visualizing spherical tensors is basically impossible. Certainly, $\vec{e}^{(1)}_{\pm 1}$ are two straigtforward linear combinations of $\widehat{x}$ and $\widehat{y}$, and $\vec{e}^{(1)}_{0} = \widehat{z}$ identically. But the use of complex coefficients for the $m = \pm 1$ elements makes visualization impossible. Beyond rank 1, it is hopeless. This is really no different from the way that it is difficult to visualize Cartesian tensors of rank $n > 1$. So, one just has to accept that visualization is difficult and that one has to build up intuition in a different way, through the mathematics.

Finally, because spherical tensors of rank j are the space $\mathbb{V}^{(j)}$, the inner product of the latter carries over, giving

$$\langle \mathcal{T}^{(j)} | \mathcal{S}^{(j)} \rangle = \underline{(\mathcal{T}^{(j)})^{*}}_{m} \underline{(\mathcal{S}^{(j)})}_{m}$$

Example 15.1: Spherical Tensors in Classical Physics: Electromagnetic Wave Polarization

There is an example of spherical tensors in classical physics, the decomposition of the EM field into circular polarization modes. Consider a plane-wave EM field propagating in the $\widehat{z} = \vec{e}_3$ direction. You know that the field can be linearly polarized, and that there one basis for this linear polarization is $\widehat{x} = \vec{e}_1$ and $\widehat{y} = \vec{e}_2$. The electric field is given by the Cartesian vector

$$\vec{E} = E_x \vec{e}_1 + E_y \vec{e}_2 \quad \underset{\vec{e}}{\longleftrightarrow} \quad \begin{bmatrix} E_x \\ E_y \\ 0 \end{bmatrix}$$

E_x and E_y must be real numbers for linear polarization so that the two polarizations are in phase and the polarization vector is time-independent.

Spherical Tensors in Classical Mechanics (cont.)

You also know that one can construct two circular polarization modes, right and left circular polarization, in which the $\vec{e}_2$ component is advanced or retarded by $\pi/2$ in phase relative to the $\vec{e}_1$ component:

$$\vec{E} = E_R \frac{1}{\sqrt{2}} (\vec{e}_1 + i\,\vec{e}_2) + E_L \frac{1}{\sqrt{2}} (\vec{e}_1 - i\,\vec{e}_2) \quad \underset{\vec{e}}{\longleftrightarrow} \quad \frac{1}{\sqrt{2}} \begin{bmatrix} E_R + E_L \\ i(E_R - E_L) \\ 0 \end{bmatrix}$$

where E_R and E_L must be real numbers to obtain circular polarization (E_x and E_y out of phase by $\pi/2$).

Spherical Tensors in Classical Mechanics (cont.)

If we allow E_x, E_y, E_R, and E_L to all be complex numbers, then we can obtain either type of polarization, or a mix of the two, in either basis. We could make circular polarization easier to write by defining basis elements

$$\vec{e}_1^{(1)} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \\ 0 \end{bmatrix} \qquad \vec{e}_{-1}^{(1)} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \\ 0 \end{bmatrix}$$

which yields

$$\vec{E} = E_R \vec{e}_1^{(1)} + E_L \vec{e}_{-1}^{(1)} \quad \xleftrightarrow[\vec{e}^{(1)}]{} \quad \begin{bmatrix} E_R \\ E_L \\ 0 \end{bmatrix}$$

The $\vec{e}_{\pm 1}^{(1)}$ are of course two of the three eigenvectors of rotations about the $\hat{z}$ direction. The last, $\vec{e}_3 = \hat{z} = \vec{e}_0^{(1)}$, is the same between the two bases. The latter version is thus a decomposition of the electric field in terms of the unit spherical tensors of rank 1. We thus have an explicit example of how a classical vector can be thought of as both a Cartesian tensor of rank 1 or a spherical tensor of rank 1.

Example 15.2: Spherical Tensors in Classical Physics: Gravitational Wave Polarization

A rank 2 example would be gravitational waves, which are propagating variations in the space-time metric, a rank 2 cartesian tensor in four dimensions under rotations and Lorentz transformations. Spatial rotations only affect the three spatial dimensions, so we can consider the space-space part of the metric to be a rank 2 tensor under spatial rotations. One can show that this rank 2 Cartesian tensor can be decomposed into spherical tensors of rank 2, 1, and 0 (one of each). There is a severe restriction on the form of the Cartesian tensor due to generic physical restrictions on the allowed form of the metric. These restrictions ensure that the rank 1 and rank 0 and $m = 0, \pm 1$ rank 2 spherical tensor components are never populated, leaving only the $m = \pm 2$ rank 2 spherical tensor components. With that, it is much simpler to think of a gravitational wave as rank 2 spherical tensor than a rank 2 Cartesian tensor: the Cartesian tensor carries along far more components than are necessary, and, even though the rank 2 spherical tensor decomposition has unnecessary components, the remaining conditions on the tensor are simplified — they are just the requirements that the $m = 0, \pm 1$ components vanish.

Connecting Cartesian Tensors and Spherical Tensors via Addition of Angular Momentum

Now that we have defined spherical tensors, we can try to qualitatively answer the question of why we chose to create them rather than just sticking with Cartesian tensors.

We already see that rank 1 Cartesian and spherical tensors, $\tau^{(1)}$ and $\mathbb{V}^{(1)}$, are the same objects, written in terms of different bases for the space. When the distinction is unimportant, we will refer to them as vectors or rank 1 tensors.

However, there is no simple correspondence between rank n Cartesian tensors and rank j spherical tensors for $n, j > 1$. Cartesian tensors of rank n have 3^n numbers in their coordinate representations, while spherical tensors of rank j have $2j + 1$ numbers. While it does hold true that, for any n, there is a j for which $3^n = 2j + 1$, we shall see later that the rotation properties of rank n Cartesian tensors are not the same as those of rank j spherical tensors with $3^n = 2j + 1$.

Spherical Tensors in Classical Mechanics (cont.)

Let us instead recall how one builds up rank n Cartesian tensors from direct products of rank 1 tensors. That relationship tells us that understanding how rank n Cartesian tensors relate to spherical tensors is a matter of asking how the space

$$\tau^{(n)} = \bigotimes_{k=1}^{n} \tau_{(k)}^{(1)} = \bigotimes_{k=1}^{n} \mathbb{V}_{(k)}^{(1)}$$

relates to the various $\mathbb{V}^{(j)}$ spaces.

Spherical Tensors in Classical Mechanics (cont.)

We will soon consider a similar problem, which is how one "adds" angular momenta. By "add", we really mean "take the direct product." For example, given a system of two particles that are each in states of orbital angular momentum $\ell = 1$, what is the angular momentum of the system as a whole? We know that the way to construct the Hilbert space for the two-particle system is to take the direct product of the individual particle Hilbert spaces. So, in this example, each particle has a Hilbert space of the form $\mathbb{V}_r \otimes \mathbb{V}_{\theta,\phi} = \mathbb{V}_r \otimes \left(\mathbb{V}^{(0)} \oplus \mathbb{V}^{(1)} \oplus \mathbb{V}^{(2)} \oplus \cdots\right)$, and, by specifying $\ell = 1$, we know that the we are picking the subspace $\mathbb{V}_r \otimes \mathbb{V}^{(1)}$. So, the joint Hilbert space of the two particles is the space

$$\left(\mathbb{V}_{(1),r} \otimes \mathbb{V}^{(1)}_{(1)}\right) \otimes \left(\mathbb{V}_{2,r} \otimes \mathbb{V}^{(1)}_{(2)}\right) = \mathbb{V}_{(1),r} \otimes \mathbb{V}_{(2),r} \otimes \left(\mathbb{V}^{(1)}_{(1)} \otimes \mathbb{V}^{(1)}_{(2)}\right)$$

where the subscript number refer to particle number (yes, different notation than we used when we originally described direct products, which is now necessary because we use $^{(j)}$ to refer to the space of states of J^2 eigenvalue j).

We will show shortly that we can decompose $\mathbb{V}^{(1)}_{(1)} \otimes \mathbb{V}^{(1)}_{(2)}$ as follows:

$$\mathbb{V}^{(1)}_{(1)} \otimes \mathbb{V}^{(1)}_{(2)} = \mathbb{V}^{(0)}_{(1)\otimes(2)} \oplus \mathbb{V}^{(1)}_{(1)\otimes(2)} \oplus \mathbb{V}^{(2)}_{(1)\otimes(2)}$$

Spherical Tensors in Classical Mechanics (cont.)

This kind of result generalizes, showing that we can decompose any direct product $\mathbb{V}^{(j_1)} \otimes \mathbb{V}^{(j_2)}$ as a direct sum of $\mathbb{V}^{(j)}$ with j running from $|j_1 - j_2|$ to $j_1 + j_2$, and you can similarly decompose a direct product of any number of $\mathbb{V}^{(j)}$. The formalism for addition of angular momenta will thus provide us a way to write the space of of Cartesian tensors of rank n as a direct sum of spaces of spherical tensors of various ranks. That is, a Cartesian tensor of rank n can be written as a sum of spherical tensors.

Finally, we see why we defined spherical tensors rather than sticking with Cartesian tensors. The above indicates that Cartesian tensors of rank n are reducible objects, meaning that the space of Cartesian tensors of rank n can be decomposed as a direct sum of spaces that are irreducible *and* invariant or closed under rotations. Quantum mechanics is always easier if we can reduce our Hilbert space to a set of irreducible closed subspaces. The Cartesian tensors of rank n are certainly invariant (closed) under rotations — a rotation does not turn a Cartesian tensor of rank n into a Cartesian tensor of a different rank — but the space is reducible, as explained above.

Another side of the same statement is that the behavior of spherical tensors under rotations is simpler than that of Cartesian tensors exactly because the spherical tensors form irreducible spaces. Irreducibility means that one cannot make the spaces smaller or simpler.

Lecture 53:
Tensor States in Quantum Mechanics
Rotations of Tensor States

Date Revised: 2009/03/07
Date Given: 2009/03/06

Tensor Particle States — Motivation

Consider a quantum-mechanical particle state $|\psi\rangle$. Since we are interested in coordinate system rotations, we must necessarily work with the position-basis representation of such a state, $\langle\vec{r}\,|\psi\rangle = \psi_q(\vec{r})$ in a coordinate system F. So far, we have considered particle states that consist of just a single number at any position in space. When we change coordinate systems from F to F' by a passive rotation transformation, or if we rotate the state itself so that the transformed state's orientation relative to F' is the same as the untransformed state's orientation relative to F, we use the formalism we developed in Section 12 to calculate

$$\psi_{q'}(\vec{r}\,') = \langle\vec{r}\,'\,|\psi\rangle = \langle\vec{r}\,|\,T^\dagger(\vec{\theta})|\psi\rangle = \langle\vec{r}\,|\exp\left(\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{L}\right)|\psi\rangle$$

$$\psi'_q(\vec{r}) = \langle\vec{r}\,|\psi'\rangle = \langle\vec{r}\,|\,T(\vec{\theta})|\psi\rangle = \langle\vec{r}\,|\exp\left(-\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{L}\right)|\psi\rangle$$

We can explicitly calculate the above by using completeness to insert $\int d^3r|\vec{r}\rangle\langle\vec{r}\,|$ between the transformation operator and $|\psi\rangle$, yielding the position-basis matrix elements of the transformation operator, which will essentially perform a Taylor expansion to rotate the wavefunction $\psi_q(\vec{r})$ around $\widehat{\theta}$ by θ.

Tensor States in Quantum Mechanics (cont.)

The point of interest here, though, is that there is a single number at each point in space. This is like a classical scalar. What if, instead, the state were specified by a vector or a tensor at each point in space? We would expect that this would be a particular coordinate representation of the vector, and that, under a rotation, there would be, in addition to the above action on the wavefunction, there would be an additional transformation of the elements of the vector at each point in space to obtain its coordinate representation in the rotated frame.

Tensor Particle States — Formal Definition

The formal way to do this is to define quantum mechanical tensor states to be states that live in the space of direct products of the kind of QM states for a particle in three spatial dimensions that we have dealt with so far and a space of classical tensors (Cartesian or spherical). That is, if $\mathbb{V}$ is a QM Hilbert space for a particle in three dimensions, then we **define** Hilbert spaces of spherical tensor states of rank j and Cartesian tensor states of rank n:

$$\mathbb{S}^{(j)} = \mathbb{V} \otimes \mathbb{V}^{(j)} \qquad \mathbb{W}^{(n)} = \mathbb{V} \otimes \tau^{(n)} \tag{15.11}$$

We have proven that $\mathbb{V}^{(j)}$ and $\tau^{(n)}$ are Hilbert spaces, so the standard properties of direct product Hilbert spaces follow: a basis for the direct product space is provided by all pairs of basis elements of the factor spaces, the inner product of the direct product space is the product of the inner products of the factor spaces, etc.

We will refer to spaces like $\mathbb{V}$ as **scalar Hilbert spaces** to distinguish them from the **tensor Hilbert spaces** of the type we are now defining.

Tensor States in Quantum Mechanics (cont.)

If $\{|v_k\rangle\}$ are a basis for $\mathbb{V}$ and we use the standard spherical unit tensors $\{|j,m\rangle\}$ for a particular coordinate system F as a basis for $\mathbb{V}^{(j)}$ and the standard Cartesian unit tensors $\{\mathcal{E}_{j_1\cdots j_n}\}$ of that coordinate system as a basis for $\tau^{(n)}$ as we defined earlier, then bases for the spherical and Cartesian tensor spaces are

$$|v^{(j)}_{k,m}\rangle = |v_k\rangle \otimes |j,m\rangle \tag{15.12}$$

$$|v_{k,j_1\cdots j_n}\rangle = |v_k\rangle \otimes |\mathcal{E}_{j_1\cdots j_n}\rangle = |v_k\rangle \otimes \bigotimes_{p=1}^{n} |\vec{e}_{j_p}\rangle \tag{15.13}$$

where we use the ket notation for the classical tensors to remind us that they are elements of a Hilbert space and, in the last expression, we recall that the unit Cartesian tensors of rank n are constructed by direct product of unit Cartesian vectors. *We emphasize that a coordinate system must be specified in choosing the* $\{|j,m\rangle\}$ *and* $\{\mathcal{E}_{j_1\cdots j_n}\}$ *because they are the unit tensors of a particular coordinate system.* This is a new phenomenon: our scalar Hilbert space states have not required the choice of a coordinate system; a coordinate system has only been necessary when we want to project onto a particular position basis $|x,y.z\rangle$ of the Hilbert space. This is an important fact: it says that the orientation information we are adding by constructing tensor states must reference a coordinate system.

Tensor States in Quantum Mechanics (cont.)

Any state in the direct product space thus has the expansion

$$|\psi^{(j)}\rangle = \sum_k \sum_{m=-j}^{j} \psi^{(j)}_{km} |v_k\rangle \otimes |j,m\rangle \qquad |\Psi^{(n)}\rangle = \sum_k \sum_{j_1\cdots j_n=1}^{3} \Psi^{(n)}_{kj_1\cdots j_n} |v_k\rangle \otimes |\mathcal{E}_{j_1\cdots j_n}\rangle$$

We designate spherical tensor states by using lower case greek letters and the $^{(j)}$ superscript and we designate Cartesian tensor states by using upper case greek letters and the $^{(n)}$ superscript. Realize that only for $j = n = 1$ can the two kinds of states be directly related. For $j \neq 1$ and $n \neq 1$, the relation will be more complex.

We obtain these expansion coefficients by projecting the state onto the appropriate basis element:

$$(\langle v_k| \otimes \langle j,m|)\,|\psi^{(j)}\rangle = \psi^{(j)}_{km} \qquad \left(\langle v_k| \otimes \langle \mathcal{E}_{j_1\cdots j_n}|\right)|\Psi^{(n)}\rangle = \Psi^{(n)}_{kj_1\cdots j_n}$$

The inner product is, as usual, the product of the inner products in the factor spaces:

$$\begin{aligned}\langle \phi^{(j)} \,|\psi^{(j)} \rangle &= \left(\sum_{k_1,m_1} \phi^{(j)*}_{k_1 m_1} \langle v_{k_1} \,| \otimes \langle j, m \,| \right) \left(\sum_{k_2,m_2} \psi^{(j)}_{k_2 m_2} |v_{k_2} \rangle \otimes |j, m_2 \rangle \right) \\ &= \sum_{k_1,k_2,m_1,m_2} \phi^{(j)*}_{k_1 m_1} \psi^{(j)}_{k_2 m_2} \langle v_{k_1} \,| v_{k_2} \rangle \langle j, m_1 \,| j, m_2 \rangle \\ &= \sum_{k_1,k_2,m_1,m_2} \phi^{(j)*}_{k_1 m_1} \psi^{(j)}_{k_2 m_2} \delta_{k_1 k_2} \delta_{m_1 m_2} = \sum_{k_1,m_1} \phi^{(j)*}_{k_1 m_1} \psi^{(j)}_{k_1 m_1}\end{aligned}$$

Similarly, for a Cartesian tensor state

$$\langle \Phi^{(n)} \,|\Psi^{(n)} \rangle = \sum_{k_1,j_1 \cdots j_n} \Phi^{(j)*}_{k_1 j_1 \cdots j_n} \Psi^{(n)}_{k_1 j_1 \cdots j_n}$$

Tensor States in Quantum Mechanics (cont.)

Finally, you can think of a tensor state as a tensor whose elements in a particular coordinate frame (its coordinate representation in that frame) consists of a set of scalar states obtained by taking the inner product with the unit tensors of that coordinate frame:

$$\underline{\left(|\psi^{(j)}\rangle\right)}_{m} = \langle j, m\,|\psi^{(j)}\rangle = \sum_k \psi^{(j)}_{km}|v_k\rangle$$

$$\underline{\left(|\Psi^{(n)}\rangle\right)}_{j_1\cdots j_n} = \langle \mathcal{E}_{j_1\cdots j_n}\,|\Psi^{(n)}\rangle = \sum_k \Psi^{(n)}_{kj_1\cdots j_n}|v_k\rangle$$

Note that these coordinate representations are frame-dependent. Hence we see how the orientation information is provided by these tensor states: the components of the coordinate representation of the state depend on the coordinate system. The scalar states themselves also depend on the coordinate system in the usual fashion for scalar states (*i.e.*, the wavefunction looks different in different coordinate systems), but it is the added orientation information provided by the different components that provide the "spin" information.

Spin-j Particles, Translation to Textbook Notation

We **define** a spin-j particle in three spatial dimensions to be a spherical tensor state of rank j because the behavior of the spherical tensor factor of the state is the same as the of an angular momentum j. Suppose $\{|n, \ell, m_\ell\rangle\}$ are a basis for the scalar Hilbert space describing the spatial behavior of the particle in terms of eigenstates of a spherically symmetric Hamiltonian and the eigenstates of L_z and L^2; then an arbitrary state may be written via the expansion

$$|\psi^{(j)}\rangle = \sum_{n,\ell,m_\ell,m} \psi^{(j)}_{n,\ell,m_\ell,m} |n, \ell, m_\ell\rangle \otimes |j, m\rangle$$

In textbooks, one usually sees the state written in one of two strange hybrid notations. In the first, one writes out the spherical tensor factor as a column matrix in the matrix representation for its $|j, m\rangle$ basis, but leaves the scalar factor as a Hilbert space state:

$$|\psi^{(j)}\rangle \xleftrightarrow[|j,m\rangle]{} \begin{bmatrix} \sum_{n,\ell,m_\ell} \psi^{(j)}_{n,\ell,m_\ell,j} |n, \ell, m_\ell\rangle \\ \sum_{n,\ell,m_\ell} \psi^{(j)}_{n,\ell,m_\ell,j-1} |n, \ell, m_\ell\rangle \\ \vdots \\ \sum_{n,\ell,m_\ell} \psi^{(j)}_{n,\ell,m_\ell,-j} |n, \ell, m_\ell\rangle \end{bmatrix} \equiv \begin{bmatrix} |\psi^{(j)}_j\rangle \\ |\psi^{(j)}_{j-1}\rangle \\ \vdots \\ |\psi^{(j)}_{-j}\rangle \end{bmatrix}$$

or in a form in which the scalar factor is projected onto the position basis:

$$|\psi^{(j)}\rangle \xleftrightarrow[|r,\theta,\phi\rangle\otimes|j,m\rangle]{} \begin{bmatrix} \sum_{n,\ell,m_\ell} \psi^{(j)}_{n,\ell,m_\ell,j}\langle r,\theta,\phi\,|n,\ell,m_\ell\rangle \\ \sum_{n,\ell,m_\ell} \psi^{(j)}_{n,\ell,m_\ell,j-1}\langle r,\theta,\phi\,|n,\ell,m_\ell\rangle \\ \vdots \\ \sum_{n,\ell,m_\ell} \psi^{(j)}_{n,\ell,m_\ell,-j}\langle r,\theta,\phi\,|n,\ell,m_\ell\rangle \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{n,\ell,m_\ell} \psi^{(j)}_{n,\ell,m_\ell,j} R_{n\ell}(r) Y_\ell^{m_\ell}(\theta,\phi) \\ \sum_{n,\ell,m_\ell} \psi^{(j)}_{n,\ell,m_\ell,j-1} R_{n\ell}(r) Y_\ell^{m_\ell}(\theta,\phi) \\ \vdots \\ \sum_{n,\ell,m_\ell} \psi^{(j)}_{n,\ell,m_\ell,-j} R_{n\ell}(r) Y_\ell^{m_\ell}(\theta,\phi) \end{bmatrix} \equiv \begin{bmatrix} \psi^{(j)}_{j}(r,\theta,\phi) \\ \psi^{(j)}_{j-1}(r,\theta,\phi) \\ \vdots \\ \psi^{(j)}_{-j}(r,\theta,\phi) \end{bmatrix}$$

These are the kinds of notation Shankar uses in Exercise 12.5.1 and Chapter 14, especially Equation (14.3.13).

Rotation Operators for Tensor Particle States, Properties under Rotation

The natural rotation operator for tensor states is obviously the tensor product of rotation operators in the factor spaces. Here we write the operators in a basis-free form:

$$\text{spherical: } U^{(j)}(\vec{\theta}) = T(\vec{\theta}) \otimes T^{(j)}(\vec{\theta}) \tag{15.14}$$

$$\text{Cartesian: } U^{(n)}(\vec{\theta}) = T(\vec{\theta}) \otimes R^{(n)}(\vec{\theta}) \tag{15.15}$$

where the $^{(j)}$ and $^{(n)}$ superscripts indicate that these operators act on rank j spherical or rank n Cartesian tensors states or tensors, as the case may be.

Tensor States in Quantum Mechanics (cont.)

The action of the above operators on spherical and Cartesian tensor states (active transformation) are as follows:

$$\text{active spherical:} \quad |\psi^{(j)\prime}\rangle = U^{(j)}(\vec{\theta})|\psi^{(j)}\rangle = U^{(j)}(\vec{\theta})\sum_{k,m}\psi^{(j)}_{km}|v_k\rangle\otimes|\vec{e}^{(j)}_m\rangle$$

$$= \sum_{km}\psi^{(j)}_{km}\left[T(\vec{\theta})|v_k\rangle\right]\otimes\left[T^{(j)}(\vec{\theta})|\vec{e}^{(j)}_m\rangle\right] \tag{15.16}$$

$$\text{active Cartesian:} \quad |\Psi^{(n)\prime}\rangle = U^{(n)}(\vec{\theta})|\Psi^{(n)}\rangle = U^{(n)}(\vec{\theta})\sum_{k,j_1,\ldots,j_n}\Psi^{(n)}_{kj_1\cdot j_n}|v_k\rangle\otimes|\mathcal{E}_{j_1\cdots j_n}\rangle$$

$$= \sum_{k,j_1,\ldots,j_n}\Psi^{(n)}_{k,j_1\cdot j_n}\left[T(\vec{\theta})|v_k\rangle\right]\otimes\left[R^{(n)}(\vec{\theta})|\mathcal{E}_{j_1\cdots j_n}\rangle\right]$$

$$= \sum_{k,j_1,\ldots,j_n}\Psi^{(n)}_{k,j_1\cdot j_n}\left[T(\vec{\theta})|v_k\rangle\right]\otimes\bigotimes_{a=1}^{n}\left[R^{(1)}(\vec{\theta})|\vec{e}_{j_a}\rangle\right] \tag{15.17}$$

which performes the desired active rotation on each factor of the direct product state. Recall that $R^{(1)}(\vec{\theta}) = T^{(1)}(\vec{\theta})$; we use different symbols only so Equations 15.14 and 15.15 can be written in a compact form.

Tensor States in Quantum Mechanics (cont.)

Of course, the above basis-free expressions are not useful calculationally, so let's write out how the matrix representations — the expansion coefficients in terms of the above bases — are related:

$$\begin{aligned}
\psi^{(j)\,\prime}_{km} &= \left(\langle v_k | \otimes \langle \vec{e}^{\,(j)}_m | \right) | \psi^{(j)\,\prime} \rangle \\
&= \sum_{pq} \left[\langle v_k | T(\vec{\theta}) | v_p \rangle \right] \left[\langle \vec{e}^{\,(j)}_m | T^{(j)}(\vec{\theta}) | \vec{e}^{\,(j)}_q \rangle \right] \psi^{(j)}_{pq} \\
&= \sum_{pq} \left[\langle v_k | T(\vec{\theta}) | v_p \rangle \right] \left[\mathbf{R}^{(j)}_{\vec{\theta}} \right]_{mq} \psi^{(j)}_{pq} \qquad (15.18) \\
\Psi^{(n)\,\prime}_{kj_1 \cdots j_n} &= \left(\langle v_k | \otimes \bigotimes_{a=1}^{n} \langle \vec{e}_{j_a} | \right) | \Psi^{(n)\,\prime} \rangle \\
&= \sum_{pq_1 \cdots q_n} \left[\langle v_k | T(\vec{\theta}) | v_p \rangle \right] \left(\prod_{a=1}^{n} \left[\langle \vec{e}_{j_a} | R^{(1)}(\vec{\theta}) | \vec{e}_{q_a} \rangle \right] \right) \Psi^{(j)}_{pq_1 \cdots q_n} \\
&= \sum_{pq_1 \cdots q_n} \left[\langle v_k | T(\vec{\theta}) | v_p \rangle \right] \left(\prod_{a=1}^{n} \left[\mathbf{R}_{\vec{\theta}} \right]_{j_a q_a} \right) \Psi^{(j)}_{pq_1 \cdots q_n} \qquad (15.19)
\end{aligned}$$

Tensor States in Quantum Mechanics (cont.)

If we are dealing with rotations, it is almost certain that we will want to use a basis for the scalar space consisting of eigenstates of L^2 and L_z (here, we are allowed to specific orbital angular momentum because we know the space of scalar states refers to the states of a particle in three spatial dimensions). That is, there will be a basis of states $\{|v_{n\ell m_\ell}\rangle = |n,\ell\rangle \otimes |\ell, m_\ell\rangle\}$ where n refers to the radial state index (quantum number) and $\{|n,\ell\rangle\}$ are states describing the radial behavior (which depend on ℓ through the centrifugal effective potential term) and $\{|\ell, m_\ell\rangle\}$ are states describing the angular dependence (the usual spherical harmonics in the position basis). This lets us write our undefined scalar space matrix elements more explicitly:

$$\begin{aligned}
\langle v_k | T(\vec{\theta}) | v_p \rangle \quad \longrightarrow \quad & \left(\langle k, \ell_1 | \otimes \langle \ell_1, m_{\ell_1} |\right) T(\vec{\theta}) \left(|p, \ell_2\rangle \otimes |\ell_2, m_{\ell_2}\rangle\right) \\
& = \langle k, \ell_1 | p, \ell_2 \rangle \langle \ell_1, m_{\ell_1} | T(\vec{\theta}) | \ell_2, m_{\ell_2} \rangle \\
& = \delta_{kp}\delta_{\ell_1\ell_2} \left[\mathbf{R}^{(\ell_1)}_{\vec{\theta}}\right]_{m_{\ell_1} m_{\ell_2}}
\end{aligned}$$

So, we have

$$\psi^{(j)\prime}_{k\ell m_{\ell,1} m} = \sum_{m_{\ell,2},q} \left[\mathbf{R}^{(\ell_1)}_{\vec{\theta}}\right]_{m_{\ell,1} m_{\ell,2}} \left[\mathbf{R}^{(j)}_{\vec{\theta}}\right]_{mq} \psi^{(j)}_{k\ell m_{\ell,2} q} \tag{15.20}$$

$$\Psi^{(n)\prime}_{k\ell m_{\ell,1} j_1 \cdots j_n} = \sum_{m_{\ell,2} q_1 \cdots q_n} \left[\mathbf{R}^{(\ell_1)}_{\vec{\theta}}\right]_{m_{\ell,1} m_{\ell,2}} \left(\prod_{a=1}^{n} \left[\mathbf{R}_{\vec{\theta}}\right]_{j_a q_a}\right) \Psi^{(j)}_{p q_1 \cdots q_n} \tag{15.21}$$

Here we see very explicitly the utility of using spherical tensor states instead of Cartesian tensor states. With spherical tensor states, the rotation of the orbital portion and of the spin portion are identical in form, using the matrix $\mathbf{R}^{(j)}_{\vec{\theta}}$. With Cartesian tensor states, the rotation of the orbital and spin components are different. And, of course, there are a lot more indices on the Cartesian component. This complexity will be reflected below in the way that the angular momentum operators act on the two kinds of states.

Lecture 54:
Angular Momentum for Tensor States
Addition of Angular Momentum

Date Revised: 2009/03/1
Date Given: 2009/03/09

Angular Momentum Operators for Tensor Particle States

We have written the rotation operators and properties of tensor states in a direct product form, rotating the scalar state and the classical tensor pieces separately. In many cases, this is not useful because the Hamiltonian depends on the total angular momentum, which has a scalar piece (orbital angular momentum) and a tensor piece (spin angular momentum). We want to see whether there is a way to decompose our tensor states into states of well-defined total angular momentum.

Tensor States in Quantum Mechanics (cont.)

Let's write the above motivation mathematically. Our rotation operators are

$$\text{spherical}: \quad U^{(j)}(\theta) = T(\vec{\theta}) \otimes T^{(j)}(\vec{\theta}) = \exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot\vec{L}\right) \otimes \exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot\vec{J}^{(j)}\right)$$

$$\begin{aligned}\text{Cartesian}: \quad U^{(n)}(\theta) &= T(\vec{\theta}) \otimes R^{(n)}(\vec{\theta}) = T(\vec{\theta}) \otimes \bigotimes_{a=1}^{n} R^{(1)}_{(a)}(\vec{\theta}) \\ &= \exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot\vec{L}\right) \otimes \bigotimes_{a=1}^{n} \exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot i\,\hbar\,\vec{M}^{(1)}_{(a)}\right) \\ &= \exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot\vec{L}\right) \otimes \bigotimes_{a=1}^{n} \exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot\vec{J}^{(1)}_{(a)}\right)\end{aligned}$$

where $\vec{L}$ is the appropriate angular momentum operator for the scalar Hilbert space (it generates rotations of scalar states), $\vec{J}^{(j)}$ is the angular momentum operator for the $\mathbb{V}^{(j)}$ space, and $\vec{M}^{(1)}_{(a)}$ is the operator corresponding to the generator of classical rotations in three spatial dimensions for the ath factor space. $\vec{J}^{(1)}$, $i\,\hbar\,\vec{M}^{(1)}$, $\vec{J}^{(1)}_{(a)}$, and $i\,\hbar\,\vec{M}^{(1)}_{(a)}$ are the same operator, though we usually use $\vec{J}^{(1)}$ when we are acting on spherical tensors and $i\,\hbar\,\vec{M}^{(1)}$ when we act on Cartesian vectors, and we only use the $_{(a)}$ subscript when we need to specify action in a particular factor space.

Tensor States in Quantum Mechanics (cont.)

So, we see that a rotation operation consists of simultaneous rotation operations in a set of factor spaces. Because the total angular momentum of the system should be tied to the rotation properties of the state as a whole, rather than that of the factor states, we want to see whether we can write the above as a single rotation operator. We can see that this should be possible by recalling that the product of the exponentials of the generators in the factor spaces is the exponential of the sum of the generators in the direct product space because the generators in the different factor spaces commute, which we proved in connection with Equation 15.8, the generator operator for rotations of Cartesian tensors. That is, we may write

$$\begin{aligned}\text{spherical}: \quad U^{(j)}(\theta) &= \exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot\vec{L}\right)\otimes\exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot\vec{J}^{(j)}\right)\\ &= \exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot\left[\vec{L}\otimes I^{(j)}+I\otimes\vec{J}^{(j)}\right]\right) \qquad (15.22)\end{aligned}$$

$$\begin{aligned}\text{Cartesian}: \quad U^{(n)}(\theta) &= \exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot\vec{L}\right)\otimes\bigotimes_{a=1}^{n}\exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot\vec{J}^{(1)}_{(a)}\right)\\ &= \exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot\left[\vec{L}\otimes\bigotimes_{a=1}^{n}I^{(1)}_{(a)}+I\otimes\left(\sum_{k=1}^{n}\vec{J}^{(1)}_{(k)}\otimes\bigotimes_{a\neq k}^{n}I^{(1)}_{(a)}\right)\right]\right) \qquad (15.23)\end{aligned}$$

Tensor States in Quantum Mechanics (cont.)

Based on the above exponential form for the rotation operator in the direct product space, we are led to define the **total angular momentum operator** $\vec{\mathcal{J}}$ by

$$\text{spherical:} \quad \vec{\mathcal{J}}^{(j)} = \vec{L} \otimes I^{(j)} + I \otimes \vec{J}^{(j)} \tag{15.24}$$

$$\text{Cartesian:} \quad \vec{\mathcal{J}}^{(n)} = \vec{L} \otimes \bigotimes_{a=1}^{n} I_{(a)}^{(1)} + I \otimes \left(\sum_{k=1}^{n} \vec{J}_{(k)}^{(1)} \otimes \bigotimes_{a \neq k}^{n} I_{(a)}^{(1)} \right) \tag{15.25}$$

That is, $\vec{\mathcal{J}}^{(j)}$ generates rotations of spherical tensor states of rank j and $\vec{\mathcal{J}}^{(n)}$ generates rotations of Cartesian tensor states of rank n. We will naturally want to find the eigenvectors and eigenvalues of $\mathcal{J}_z$ and $\mathcal{J}^2$ for each kind of tensor. We recognize that the two problems are essentially the same in that they consist of finding the eigevectors and eigenvalues of the sum of two or more angular momentum operators. Thus, we will need to turn to the problem of addition of angular momentum in order to sort out the rotation properties of tensor states in terms of their total angular momentum.

Section 16

Addition of Angular Momenta

Addition of Angular Momentum – States

Overview

We will do two things in this section:

- We will formally show that if one "adds" two angular momenta j_1 and j_2 by taking the direct product of their angular momentum spaces $V^{(j_1)}$ and $V^{(j_2)}$, $j_1 \geq j_2$, then one obtains a direct sum of all angular momentum spaces between $V^{(j_1+j_2)}$ and $V^{(j_1-j_2)}$, inclusive:

$$V^{(j_1)} \otimes V^{(j_2)} = V^{(j_1+j_2)} \oplus V^{(j_1+j_2-1)} \oplus \ldots \oplus V^{(j_1-j_2+1)} \oplus V^{(j_1-j_2)}$$

- We will determine the generic form for the expansion coefficients needed to write elements in the direct sum space in terms of the the direct products of the basis elements in the factor spaces; that is, we will figure out how to write the obvious basis kets $\{|j, m\rangle\}$ of the direct sum space in terms of the product space basis kets $\{|j_1, m_1\rangle \otimes |j_2, m_2\rangle\}$. These expansion coefficients are called the Clebsch-Gordan coefficients.

Motivation for "Addition" of Angular Momentum

Let's first understand what we mean by "addition" of angular momentum. We have already seen a need for this in our demonstration that, when we try to add "spin" information to a particle state by considering direct products of a scalar Hilbert space $\mathbb{V}$ and a classical tensor Hilbert space, we find that the generator of rotations for the product space is

$$\text{spherical:} \quad \vec{\mathcal{J}}^{(j)} = \vec{L} \otimes I^{(j)} + I \otimes \vec{J}^{(j)}$$

$$\text{Cartesian:} \quad \vec{\mathcal{J}}^{(n)} = \vec{L} \otimes \bigotimes_{a=1}^{n} I^{(1)}_{(a)} + I \otimes \left(\sum_{k=1}^{n} \vec{J}^{(1)}_{(k)} \otimes \bigotimes_{a \neq k}^{n} I^{(1)}_{(a)} \right)$$

We have so far written states in this product space in terms of the eigenstates of L^2, L_z, $[J^2]^{(j)}$, and $J_z^{(j)}$. But, since it is the sum operator $\vec{\mathcal{J}}$ that generates rotations in the direct product space, it makes sense to want to consider a basis of eigenstates of $\mathcal{J}^2$ and $\mathcal{J}_z$. We can see how this would be useful physically, as, in many cases, we are interested in the total angular momentum of the system — summing together the orbital and spin contributions — rather than the separate pieces.

Addition of Angular Momentum – States (cont.)

Another milieu in which we would be interested in adding angular momenta is when we form multiparticle systems. For example, neglecting spin, suppose we want to consider the orbital angular momentum of a system of two particles. Clearly, if $\vec{L}_{(1)}$ and $\vec{L}_{(2)}$ are the angular momentum operators for particles 1 and 2 acting in (scalar) Hilbert spaces $\mathbb{V}_{(1)}$ and $\mathbb{V}_{(2)}$, then the generator of rotations in the product space is

$$\vec{\mathcal{J}} = \vec{L}_{(1)} \otimes I_{(2)} + I_{(1)} \otimes \vec{L}_{(2)}$$

Another example would be to consider the combined spin of a two-particle system consisting of spins j_1 and j_2:

$$\vec{\mathcal{J}} = \vec{J}^{(j_1)} \otimes I^{(j_2)} + I^{(j_1)} \otimes \vec{J}^{(j_2)}$$

Thanks to the generic properties of angular momentum operators, all three of these examples are essentially the same problem: taking the product of two angular momentum spaces. So we will find a generic solution to the problem.

Formal Decomposition of the State Space

So, let's consider two angular momentum operators, $\vec{J}^{(j_1)}$ and $\vec{J}^{(j_2)}$, acting in factor spaces $V^{(j_1)}$ and $V^{(j_2)}$. Note that it is completely generic to use the $V^{(j)}$ spaces we found in Section 14.6, regardless of whether we are considering orbital or spin angular momentum or integral or half-integral j values, because the eigenstates for any angular momentum operator can be written in that form. We also drop 3 the calligraphic font for the total angular momentum because we do not care about the distinction between the total and individual angular momenta now. So we consider

$$\vec{J} = \vec{J}^{(j_1)} \otimes I^{(j_2)} + I^{(j_1)} \otimes \vec{J}^{(j_2)} \qquad \mathbb{V} = \mathbb{V}^{(j_1)} \otimes \mathbb{V}^{(j_2)}$$

The basis states for the factor spaces are $\{|j_1, m_1\rangle\}$ and $\{|j_2, m_2\rangle\}$. The obvious basis for the product space is $\{|j_1, m_1\rangle \otimes |j_2, m_2\rangle\}$. This is called the uncoupled basis or uncoupled representation because we consider the angular momentum state of each factor separately.

Addition of Angular Momentum – States (cont.)

The first thing we recognize is that $\vec{J}$ satisfies the standard angular momentum commutation relations,

$$[J_a, J_b] = i\,\hbar\,\epsilon_{abc} J_c$$

simply because $\vec{J}$ is the sum of two commuting terms that separately satisfy the above. This immediately tells us that our standard angular momentum formalism is valid for $\vec{J}$; that is:

1. J^2 is allowed to have eigenvalues of the form $\alpha = \hbar^2 j\,(j+1)$, j any positive integer or half-integer
2. J_z is allowed to have eigenvalues $m = j,\, j-1,\, \ldots,\, -j+1,\, -j$
3. The space of states on which $\vec{J}$ operates has a basis $\{|j, m\rangle\}$ of simultaneous eigenstates of J^2 and J_z labeled by their eigenvalues j and m. All we need to determine now is which values of j are in use and how the $\{|j, m\rangle\}$ are related to the $\{|j_1, m_1\rangle \otimes |j_2, m_2\rangle\}$.

Addition of Angular Momentum – States (cont.)

The second thing we notice is that, because J_z commutes with each of its contributing terms, $J_z^{(j_1)} \otimes I^{(j_2)}$ and $I^{(j_1)} \otimes J_z^{(j_2)}$, eigenstates of the factor operators are eigenstates of J_z. From the form of J_z, the J_z eigenvalue of a product state $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$ is just the sum of the eigenvalues of the factor states; $m = m_1 + m_2$. This immediately tells us what values of m are accessible: $m = j_1 + j_2$ to $m = -(j_1 + j_2)$. So, we already know that the product space must contain $\mathbb{V}^{(j_1+j_2)}$; that is, $\mathbb{V}^{(j_1+j_2)}$ is a subspace of $\mathbb{V}$. Moreover, just by counting states, we know that $\mathbb{V}^{(j_1+j_2)}$ cannot be all of $\mathbb{V}$: $\mathbb{V}$ has $(2j_1+1)(2j_2+1)$ states, while $\mathbb{V}^{(j_1+j_2)}$ only has $2j_1 + 2j_2 + 1$ states; there are $4j_1 j_2$ states to be identified. Those other states must live in other subspaces of $\mathbb{V}$. Finally, because of the angular momentum structure of $\vec{J}$, those subspaces must be $\mathbb{V}^{(j)}$ for some values of j to be determined. That is, we already know

$$\mathbb{V}^{(j_1)} \otimes \mathbb{V}^{(j_2)} = \mathbb{V}^{(j_1+j_2)} \oplus \sum_{j \ \ tbd} \mathbb{V}^{(j)}$$

where j *tbd* means the j values are to be determined.

Addition of Angular Momentum – States (cont.)

The obvious next question is – what other j values are subspaces of $\mathbb{V}$? We can see this by walking down the possible values of J_z and counting states. Let's list the number of possible uncoupled basis states that could yield states of a particular J_z eigenvalue m (the latter we will call the coupled basis). Certainly, $m = m_1 + m_2$ is required; then we have

J_z	contributing (m_1, m_2) values
$j_1 + j_2$	(j_1, j_2)
$j_1 + j_2 - 1$	$(j_1, j_2 - 1),\ (j_1 - 1, j_2)$
$j_1 + j_2 - 2$	$(j_1, j_2 - 2),\ (j_1 - 1, j_2 - 1),\ (j_1 - 2, j_2)$
$\vdots$	$\vdots$
$j_1 + j_2 - n$	$(j_1, j_2 - n),\ (j_1 - 1, j_2 - (n-1)), \ldots, (j_1 - (n-1), j_2 - 1),\ (j_1 - n, j_2)$
$\vdots$	$\vdots$
$j_1 - j_2$	$(j_1, -j_2),\ (j_1 - 1, -j_2 + 1), \ldots, (j_1 - 2j_2 + 1, j_2 - 1),\ (j_1 - 2j_2, j_2)$
$j_1 - j_2 - 1$	$(j_1 - 1, -j_2),\ (j_1 - 2, -j_2 + 1), \ldots, (j_1 - 1 - 2j_2 + 1, j_2 - 1),\ (j_1 - 1 - 2j_2, j_2)$
$\vdots$	$\vdots$

Addition of Angular Momentum – States (cont.)

We note the following:

- There is a clear pattern in the first $2j_2 + 1$ states in the coupled basis: there are $n + 1$ states that can yield $m = j_1 + j_2 - n$ for $n = 0, \ldots, 2j_2$. $\mathbb{V}^{(j_1+j_2)}$ can only provide one of these for each value of m, so, as we noted, additional spaces are needed in the direct sum. For example, for $n = 1$, $m = j_1 + j_2 - 1$, there are two uncoupled basis states, so there must be two states with $m = j_1 + j_2 - 1$ in the coupled basis. Only states in $\mathbb{V}^{(j)}$ for $j \geq j_1 + j_2 - 1$ can yield such a state. But if $j \geq j_1 + j_2$, then $\mathbb{V}^{(j)}$ will also yield states with $m \geq j_1 + j_2$, and we don't need any such states — we know there is only one state with $m = j_1 + j_2$ in either the coupled or uncoupled basis. So there is only one choice for the space to add to the direct sum, and that is $\mathbb{V}^{(j_1+j_2-1)}$. A similar argument holds for the remaining $n = 2, 3, \ldots 2j_2$. So we successively add $\mathbb{V}^{(j_1+j_2-2)}$, $\mathbb{V}^{(j_1+j_2-3)}$, …, $\mathbb{V}^{(j_1+j_2-2j_2)} = \mathbb{V}^{(j_1-j_2)}$. That is,

$$\mathbb{V}^{(j_1)} \otimes \mathbb{V}^{(j_2)} = \mathbb{V}^{(j_1+j_2)} \oplus \mathbb{V}^{(j_1+j_2-1)} \oplus \ldots \oplus \mathbb{V}^{(j_1-j_2+1)} \oplus \mathbb{V}^{(j_1-j_2)} \oplus \ldots$$

Addition of Angular Momentum – States (cont.)

- Moreover, we can conclude that the direct sum terminates at $\mathbb{V}^{(j_1-j_2)}$, that the $\cdots$ at the end are unnecessary, by just counting states. The space $\mathbb{V}^{(j)}$ contributes $2j+1$ states. The total number of states we have accounted for so far is

$$\begin{aligned} N &= \sum_{j=j_1-j_2}^{j_1+j_2} 2j+1 = \sum_{j=0}^{j_1+j_2} (2j+1) - \sum_{j=0}^{j_1-j_2-1} (2j+1) \\ &= (j_1+j_2)(j_1+j_2+1) + (j_1+j_2+1) \\ &\quad - (j_1-j_2-1)(j_1-j_2) - (j_1-j_2) \\ &= 4j_1j_2 + 2j_1 + 2j_2 + 1 = (2j_1+1)(2j_2+1) \end{aligned}$$

where we have used $\sum_{n=0}^{N} = N(N+1)/2$. So, we have accounted for all the uncoupled basis states. Therefore, we have $2j_2+1$ terms in the direct sum,

$$\mathbb{V}^{(j_1)} \otimes \mathbb{V}^{(j_2)} = \mathbb{V}^{(j_1+j_2)} \oplus \mathbb{V}^{(j_1+j_2-1)} \oplus \ldots \oplus \mathbb{V}^{(j_1-j_2+1)} \oplus \mathbb{V}^{(j_1-j_2)}$$

Addition of Angular Momentum – States (cont.)

- For completness, let's explain which states contribute to m for $m < j_1 - j_2$. For $m \leq -(j_1 - j_2)$, the situation is a mirror of what we have done so far, so that's trivial. For $0 \leq m < j_1 - j_2$, one can't begin with $m_1 = j_1$ (as we do for $m \geq j_1 - j_2$) because m_2 cannot be large and negative enough to yield $m < j_1 - j_2$; it would require $m_2 < -j_2$, which is not possible. So, to obtain a coupled state with $m = j_1 - j_2 - n$, one begins with $(m_1 = j_1 - n, m_2 = -j_2)$ and ends with $(m_1 = j_1 - 2j_2 - n, m_2 = j_2)$; there are $2j_2 + 1$ states that contribute to each m. That exactly matches the number of terms in the direct sum, and each term of the direct sum contributes one state, so the counting is correct. This works all the way down to $m = 0$ or $m = 1/2$ (depending on whether exactly one of j_1 and j_2 is half-integer or not), which corresponds to $n = j_1 - j_2$ or $n = j_1 - j_2 - 1/2$. Then, of course, for $-(j_1 - j_2) + 1 \leq m < 0$, we just mirror $0 < m \leq j_1 - j_2 - 1$.
- Finally, we comment that the states in the coupled representation are linear combinations of the states in the uncoupled basis. For example, for $m = j_1 + j_2 - 1$, there are two coupled basis states to obtain, $|j = j_1 + j_2, m = j_1 + j_2 - 1\rangle$ and $|j = j_1 + j_2 - 1, m = j_1 + j_2 - 1\rangle$, and two uncoupled basis states that contribute, $|j_1, m_1 = j_1\rangle \otimes |j_2, m_2 = j_2 - 1\rangle$ and $|j_1, m_1 = j_1 - 1\rangle \otimes |j_2, m_2 = j_2\rangle$. The coupled basis states are orthogonal combinations of the two uncoupled basis states, as we will show below. In general, the n coupled basis states for some m are n orthogonal linear combinations of the uncoupled basis states that yield that m. We shall show this explicitly below.

Addition of Angular Momentum – States (cont.)

Some examples:

- Combination of orbital and spin angular momentum in the 3-dimensional SHO. Recall that the eigenvalues and allowed ℓ values for the 3D SHO are

$$E_{n,\ell,m} = \left(n + \frac{3}{2}\right)\hbar\omega \qquad \ell = n, n-2, \ldots, 1 \text{ or } 0$$

  So ℓ can take on integer values. Consider the $n = 1$ state, which allows $\ell = 1$ only. Suppose that the particle in the SHO potential is an electron with spin $1/2$. We have $j_1 = 1$ and $j_2 = 1/2$, so the allowed coupled basis j values are $j = 3/2$ and $j = 1/2$. There are $3 \times 2 = 6$ uncoupled basis states from $j_1 = 1$ and $j_2 = 1/2$, and there are $4 + 2 = 6$ coupled basis states from $j = 3/2$ and $j = 1/2$.

Addition of Angular Momentum – States (cont.)

- Allowed total spin for the proton. The proton consists of three spin-1/2 quarks. The spin combinations, combined with possible additional orbital angular momentum of the three quarks, will add together to yield the total spin of the proton. Let's first combine the spins of the three quarks. We will begin to use the notation j to replace the more cumbersome $\mathbb{V}^{(j)}$ where it will cause no confusion. We have

$$\begin{aligned}\frac{1}{2}\otimes\frac{1}{2}\otimes\frac{1}{2} &= \frac{1}{2}\otimes(1\oplus 0) = \left(\frac{1}{2}\otimes 1\right)\oplus\left(\frac{1}{2}\otimes 0\right)\\ &= \left(\frac{3}{2}\oplus\frac{1}{2}\right)\oplus\frac{1}{2} = \frac{3}{2}\oplus\frac{1}{2}\oplus\frac{1}{2}\end{aligned}$$

We see two interesting things in the above. First, we note the "distributivity" property of direct products over direct sums. This can be checked by considering the above expression for basis elements. Second, we see that two spin 1/2 spaces appear in the end. That is, the resulting Hilbert space has two subspaces that both look like $\mathbb{V}^{(1/2)}$. There is nothing wrong with that, we just have to be careful to explain which subspace we are talking about when we refer to a spin-1/2 subspace of $\mathbb{V}$.

Addition of Angular Momentum – States (cont.)

If we now add in orbital angular momentum, we are assured that the resulting Hilbert space will be a direct sum of half-integer-spin subspaces. Whether there is a spin-1/2 subspace will depend on what orbital angular momentum we add in. Obviously, adding in $\ell = 0$ will change nothing, adding $\ell = 1$ will provide access to spin 5/2 but still allow spin 1/2, etc.

We note that the different subspaces correspond to different particles. Let's consider $\ell = 0$, so what we have listed above is all that we have. One of the spin-1/2 subspaces is the proton. The spin-3/2 subspace is the Δ^+, which has the same quark content but higher spin. The other spin-1/2 subspace is (I believe) disallowed by the Pauli exclusion principle on quark *color*, not spin because the quark content of these particles is up, up, down, so two are identical.

Another example is the combination of the up, down, and strange quarks, which yield the spin-1/2 $\Sigma^0(1193)$ and $\Lambda(1116)$ particles and the spin-3/2 $\Sigma^0(1384)$ particle (the numbers in parentheses indicate the mass in MeV). Here, all three subspaces correspond to real particles because there is no Pauli exclusion restriction.

States with $\ell \neq 0$ may manifest as stable particles or as short-lived resonances.

Addition of Angular Momentum – States (cont.)

Clebsch-Gordan Coefficients

In the above, we have determined the basic structure of the product space formed by "adding" two angular momenta. Now, let's figure out in detail how to transform from the uncoupled basis to the coupled basis; essentially, how to write the natural basis of the direct sum space in terms of the natural basis of the direct product space.

The generic form for the expansion is, obviously,

$$|j,m\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} \left(|j_1,m_1\rangle \otimes |j_2,m_2\rangle\right)\left(\langle j_1,m_1| \otimes \langle j_2,m_2|\right)|j,m\rangle$$

where the expansion coefficients are called the Clebsch-Gordan (CG) coefficients. We define

$$C^{(j_1,j_2)}_{m_1 m_2 j m} = \left(\langle j_1,m_1| \otimes \langle j_2,m_2|\right)|j,m\rangle$$

How do we calculate the CG coefficients? We simply start from the top and work down, making some reasonable choices for arbitrary phase conventions along the way.

Addition of Angular Momentum – States (cont.)

In detail:

1. First, we know only one uncoupled basis state corresponds to the coupled basis state of maximum m:

$$|j = j_1 + j_2, m = j_1 + j_2\rangle = |j_1, j_1\rangle \otimes |j_2, j_2\rangle$$

We follow the Condon-Shortley convention in setting the phase factor to 1 for simplicity. There is a corresponding relation for $|j = j_1 + j_2, m = -(j_1 + j_2)\rangle$, and it turns out that setting the phase factor there to 1 also is consistent with what one would get via lowering operators.

Addition of Angular Momentum – States (cont.)

2. Second, we can obtain all the states $|j = j_1 + j_2, m\rangle$ simply by acting with the total angular momentum lowering operator J_-:

$$\begin{aligned}
&J_- |j = j_1 + j_2, m = j_1 + j_2\rangle \\
&= \left(J_-^{(j_1)} \otimes I^{(j_2)} + I^{(j_1)} \otimes J_-^{(j_2)}\right) (|j_1, j_1\rangle \otimes |j_2, j_2\rangle) \\
&= \hbar\sqrt{(j_1 + j_1)(j_1 - j_1 + 1)}\ |j_1, j_1 - 1\rangle \otimes |j_2, j_2\rangle \\
&\quad + \hbar\sqrt{(j_2 + j_2)(j_2 - j_2 + 1)}\ |j_1, j_1\rangle \otimes |j_2, j_2 - 1\rangle \\
&= \hbar\sqrt{2j_1}\ |j_1, j_1 - 1\rangle \otimes |j_2, j_2\rangle + \hbar\sqrt{2j_2}\ |j_1, j_1\rangle \otimes |j_2, j_2 - 1\rangle
\end{aligned}$$

We also expect, based on the fact that J_- is a lowering operator, that

$$\begin{aligned}
&J_- |j = j_1 + j_2, m = j_1 + j_2\rangle \\
&= \hbar\sqrt{(j_1 + j_2 + j_1 + j_2)(j_1 + j_2 - (j_1 + j_2) + 1)}\ |j = j_1 + j_2, m = j_1 + j_2 - 1\rangle \\
&= \hbar\sqrt{2(j_1 + j_2)}|j = j_1 + j_2, m = j_1 + j_2 - 1\rangle
\end{aligned}$$

Combining the two, we have

$$|j = j_1 + j_2, m = j_1 + j_2 - 1\rangle = \frac{\sqrt{j_2}\ |j_1, j_1\rangle \otimes |j_2, j_2 - 1\rangle + \sqrt{j_1}\ |j_1, j_1\rangle \otimes |j_2 - 1, j_2\rangle}{\sqrt{j_1 + j_2}}$$

Continuing downward is rather tedious but otherwise straightforward.

Addition of Angular Momentum – States (cont.)

3. To obtain the top of each ladder for $j_1 - j_2 \leq j < j_1 + j_2$, we simply require orthogonality of the top of the ladder with the higher j states already calculated, along with requiring real coefficients. For example, we find $|j = j_1 + j_2 - 1, m = j_1 + j_2 - 1\rangle$ by requiring it be a linear combination of the same uncoupled states as $|j = j_1 + j_2, m = j_1 + j_2 - 1\rangle$, but requiring the two be orthogonal and that the new state have real Clebsch-Gordan coefficients.

With that, we are able to obtain all the Clebsch-Gordan coefficients for a given problem. You can find CG coefficients for some low j combinations in the Particle Data Book at

http://pdg.lbl.gov/2007/reviews/clebrpp.pdf

The Particle Data Book is produced by the Particle Data Group at LBNL (http://pdg.lbl.gov/).

Addition of Angular Momentum – States (cont.)

General Properties of Clebsch-Gordan Coefficients

- A CG coefficient is nonvanishing only if $j_1 - j_2 \leq j \leq j_1 + j_2$

$$(\langle j_1, m_1 | \otimes \langle j_2, m_2 |) \neq 0 \quad \text{only for} \quad j_1 - j_2 \leq j \leq j_1 + j_2$$

- A CG coefficient is nonvanishing only if $m = m_1 + m_2$.
- By convention, CG coefficients are always real.
- By convention,

$$(\langle j_1, j_1 | \otimes \langle j_2, j - j_1 |) |j, j\rangle \geq 0 \qquad \text{for any} \quad j_1 - j_2 \leq j \leq j_1 + j_2$$

  This fixes the sign of the top state for each j.
- The symmetry properties under $m_1, m_2, m \longrightarrow -m_1, -m_2, -m$ imply

$$(\langle j_1, -m_1 | \otimes \langle j_2, -m_2 |) |j, -m\rangle = (-1)^{j_1+j_2-j} (\langle j_1, m_1 | \otimes \langle j_2, m_2 |) |j, m\rangle$$

- The Clebsch-Gordan coefficients define a transformation from one orthonormal basis to another, so define a unitary transformation and the elements themselves form a unitary matrix. Because all the coefficients are real, the matrix is in fact orthogonal. It is just like any unitary transformation from one basis to another.

Examples

- **Addition of two spin-1/2 angular momenta**
  Clearly, $1/2 \otimes 1/2 = 1 \oplus 0$. Let's work out the relations between the states using the above formalism. First, the top and bottom state $|j = 1, m = \pm 1\rangle$ are obviously given by

$$|j = 1, m = 1\rangle = \left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle$$
$$|j = 1, m = -1\rangle = \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle$$

  We use the formula given above for the action of the lowering operator on the top state to obtain

$$|j = 1, m = 0\rangle = \frac{1}{\sqrt{2}}\left(\left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle + \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle\right)$$

  Finally, we use the orthogonality requirement and the normalization, realness, sign-fixing conventions to obtain

$$|j = 0, m = 0\rangle = \frac{1}{\sqrt{2}}\left(\left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle - \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle\right)$$

Addition of Angular Momentum – States (cont.)

For the sake of completeness, let's invert the above. It is trivial for the $|j = 1, m = \pm 1\rangle$ states, but the other two are not:

$$\left|\frac{1}{2}, \frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \frac{1}{\sqrt{2}}\left(|j = 1, m = 0\rangle + |j = 0, m = 0\rangle\right)$$

$$\left|\frac{1}{2}, -\frac{1}{2}\right\rangle \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle = \frac{1}{\sqrt{2}}\left(|j = 1, m = 0\rangle - |j = 0, m = 0\rangle\right)$$

Addition of Angular Momentum – States (cont.)

- **Addition of $\vec{L}$ and $\vec{S}$**
  A very typical situation is addition of an orbital angular momentum with a particle spin. The former is guaranteed to be integer, and the latter is 1/2 for all fundamental particles we know of. So we have

$$\ell \otimes \frac{1}{2} = \left(\ell + \frac{1}{2}\right) \oplus \left(\ell - \frac{1}{2}\right)$$

Let's construct the states explicitly. First, of course, the top and bottom states of $j = \ell + 1/2$:

$$\left| j = \ell + \frac{1}{2}, m = \ell + \frac{1}{2} \right\rangle = |\ell, \ell\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle$$
$$\left| j = \ell + \frac{1}{2}, m = -\left(\ell + \frac{1}{2}\right) \right\rangle = |\ell, -\ell\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle$$

As usual, we use the lowering operator to obtain the next highest state:

$$\left| j = \ell + \frac{1}{2}, m = \ell - \frac{1}{2} \right\rangle = \sqrt{\frac{1}{2\ell+1}} |\ell, \ell\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle + \sqrt{\frac{2\ell}{2\ell+1}} |\ell, \ell-1\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle$$

Addition of Angular Momentum – States (cont.)

Similarly,

$$\left| j = \ell + \frac{1}{2}, m = -\left(\ell - \frac{1}{2}\right) \right\rangle$$
$$= \sqrt{\frac{1}{2\ell+1}} \left|\ell, -\ell\right\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle + \sqrt{\frac{2\ell}{2\ell+1}} \left|\ell, -(\ell-1)\right\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle$$

We determine the top and bottom states of $j = \ell - 1/2$ by requiring orthogonality, normalization, realness, and the sign-fixing convention:

$$\left| j = \ell - \frac{1}{2}, m = \ell - \frac{1}{2} \right\rangle$$
$$= \sqrt{\frac{2\ell}{2\ell+1}} \left|\ell, \ell\right\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle - \sqrt{\frac{1}{2\ell+1}} \left|\ell, \ell-1\right\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle$$

$$\left| j = \ell - \frac{1}{2}, m = -\left(\ell - \frac{1}{2}\right) \right\rangle$$
$$= \sqrt{\frac{2\ell}{2\ell+1}} \left|\ell, -\ell\right\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle - \sqrt{\frac{1}{2\ell+1}} \left|\ell, -(\ell-1)\right\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle$$

Addition of Angular Momentum – States (cont.)

In general, this is where it begins to get difficult because the next states down are composed of three uncoupled basis states. But, here, because $j_2 = 1/2$, the direct sum only has two terms and so every coupled basis state is composed of only two uncoupled basis states. Let's just write it out, leaving the coefficients to be determined recursively:

$$
\begin{aligned}
&J_- \left| j = \ell + \frac{1}{2}, m = n + \frac{1}{2} \right\rangle \\
&= \left(L_- \otimes I^{(1/2)} + I^{(\ell)} \otimes S_- \right) \left(\alpha |\ell, n+1\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle + \beta |\ell, n\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle \right) \\
&\hbar \sqrt{\left(\ell + \frac{1}{2} + n + \frac{1}{2} \right) \left(\ell + \frac{1}{2} - n - \frac{1}{2} + 1 \right)} \left| j = \ell + \frac{1}{2}, m = n - \frac{1}{2} \right\rangle \\
&= \alpha \hbar \sqrt{(\ell + n + 1)(\ell - n - 1 + 1)} \, |\ell, n\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \\
&\quad + \beta \hbar \sqrt{(\ell + n)(\ell - n + 1)} \, |\ell, n-1\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle + \beta \hbar \, |\ell, n\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle
\end{aligned}
$$

which yields

$$\begin{aligned}
&\left| j = \ell + \frac{1}{2}, m = n - \frac{1}{2} \right\rangle \\
&= \left(\alpha \sqrt{\frac{\ell - n}{\ell - n + 1}} + \beta \sqrt{\frac{1}{(\ell + n + 1)(\ell - n + 1)}} \right) |\ell, n\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \\
&\quad + \beta \sqrt{\frac{\ell + n}{\ell + n + 1}} \, |\ell, n - 1\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle
\end{aligned}$$

We have α and β for $n = \ell - 1$, from these we guess the generic formulae

$$\alpha_{n=\ell-1} = \sqrt{\frac{1}{2\ell + 1}} = \sqrt{\frac{\ell - n}{2\ell + 1}} \qquad \beta_{n=\ell-1} = \sqrt{\frac{2\ell}{2\ell + 1}} = \sqrt{\frac{\ell + n + 1}{2\ell + 1}}$$

Addition of Angular Momentum – States (cont.)

Inserting these, we find

$$\left| j = \ell + \frac{1}{2}, m = n - \frac{1}{2} \right\rangle$$
$$= \sqrt{\frac{\ell - n + 1}{2\ell + 1}} \, |\ell, n\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle + \sqrt{\frac{\ell + n}{2\ell + 1}} \, |\ell, n-1\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle$$

We see that the coefficients obey the expected formulae

$$\alpha_{n \to n-1} = \sqrt{\frac{\ell - (n-1)}{2\ell + 1}} = \sqrt{\frac{\ell - n + 1}{2\ell + 1}}$$
$$\beta_{n \to n-1} = \sqrt{\frac{\ell + (n-1) + 1}{2\ell + 1}} = \sqrt{\frac{\ell + n}{2 + 1}}$$

and that the resulting state is correctly normalized. So our guesses for α_n and β_n were correct and the above result for $|j = \ell + 1/2, m = n - 1/2\rangle$ holds in general for $n = \ell, \ldots, -\ell$.

Addition of Angular Momentum – States (cont.)

Finally, we can obtain the other state of the same m by requiring orthogonality, normalization, realness, and sign-fixing:

$$\left| j = \ell - \frac{1}{2}, m = n - \frac{1}{2} \right\rangle$$
$$= \sqrt{\frac{\ell + n}{2\ell + 1}} \, |\ell, n\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle - \sqrt{\frac{\ell - n + 1}{2\ell + 1}} \, |\ell, n - 1\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle$$

Addition of Angular Momentum – States (cont.)

- **Tensors of rank n**
  We have seen that tensors of rank 0 and 1 (scalars and vectors) correspond to angular momentum $j = 0$ and $j = 1$. Recalling that these spaces were called $\tau^{(0)}$ and $\tau^{(1)}$, and that the space of states of angular momentum j is denoted by $\mathbb{V}^{(j)}$, we have $\tau^{(0)} = \mathbb{V}^{(0)}$ and $\tau^{(1)} = \mathbb{V}^{(1)}$. Does the same correspondence hold for tensors of arbitrary rank n? No, which we can see by just counting basis elements. Tensors of rank n are a 3^n-dimensional space because there are 3^n basis elements (*e.g.*, the unit tensors). The set of states of angular momentum j is $2j+1$-dimensional because there are $2j+1$ basis elements, the $\{|j, m\rangle\}$. The two dimensionalities only coincide for $n = j = 0$ and $n = j = 1$.

  However, recall that

$$\tau^{(n)} = \bigotimes_{k=1}^{n} \tau^{(1)} = \bigotimes_{k=1}^{n} \mathbb{V}^{(1)}$$

Addition of Angular Momentum – States (cont.)

That is, the space of tensors of rank n is the direct product of n spaces of angular momentum $j = 1$, which means that it looks like the space obtained by "adding angular momentum" for n angular momenta, all with $j = 1$. The result will be of course be a direct sum space of $\mathbb{V}^{(j)}$ with $j = 0$ to $j = n$; but the trick is to see how many copies of each $\mathbb{V}^{(j)}$ we obtain. We can derive it by induction. Suppose

$$\tau^{(n)} = \sum_{j=1}^{n} C_j^{(n)} \mathbb{V}^{(j)}$$

where $C_j^{(n)}$ indicates the number of copies of $\mathbb{V}^{(j)}$ in $\tau^{(n)}$. Then it should be clear that the formula for $\tau^{(n+1)}$ is

$$\begin{aligned}
\tau^{(n+1)} &= \mathbb{V}^{(1)} \otimes \sum_{j=0}^{n} C_j^{(n)} \mathbb{V}^{(j)} \\
&= \sum_{j=1}^{n} C_j^{(n)} \left(\mathbb{V}^{(j+1)} \oplus \mathbb{V}^{(j)} \oplus \mathbb{V}^{(j-1)} \right) \oplus C_0^{(n)} \mathbb{V}^{(1)} \\
&= C_n^{(n)} \mathbb{V}^{(n+1)} + \left(C_n^{(n)} + C_{n-1}^{(n)} \right) \mathbb{V}^{(n)} \\
&\quad \oplus \sum_{j=1}^{n-1} \left(C_{j+1}^{(n)} + C_j^{(n)} + C_{j-1}^{(n)} \right) \mathbb{V}^{(j)} \oplus C_1^{(n)} \mathbb{V}^{(0)}
\end{aligned}$$

Addition of Angular Momentum – States (cont.)

Summarizing,

$$C_{n+1}^{(n+1)} = C_n^{(n)} \qquad C_n^{(n+1)} = C_n^{(n)} + C_{n-1}^{(n)} \qquad C_j^{(n+1)} = C_{j+1}^{(n)} + C_j^{(n)} + C_{j-1}^{(n)}$$
$$C_0^{(n+1)} = C_1^{(n)}$$

We must specify "initial conditions" for the recursion using $n-2$, which is

$$\tau^{(2)} = \mathbb{V}^{(1)} \otimes \mathbb{V}^{(1)} = \mathbb{V}^{(2)} \oplus \mathbb{V}^{(1)} \oplus \mathbb{V}^{(0)}$$

So,

$$C_3^{(3)} = 1 \quad C_2^{(3)} = 2 \quad C_1^{(3)} = 3 \quad C_0^{(3)} = 1$$
$$C_4^{(4)} = 1 \quad C_3^{(4)} = 3 \quad C_2^{(4)} = 6 \quad C_1^{(4)} = 6 \quad C_0^{(4)} = 3$$
$$C_5^{(5)} = 1 \quad C_4^{(5)} = 4 \quad C_3^{(5)} = 10 \quad C_2^{(5)} = 15 \quad C_1^{(5)} = 15 \quad C_0^{(5)} = 6$$

I have been unable to find a generic closed form expression (though there may be one) for any of the above except for the following:

$$C_n^{(n)} = 1 \qquad C_{n-1}^{(n)} = n-1$$

Addition of Angular Momentum – States (cont.)

Given the above, it should be clear that one can do a generalized Clebsch-Gordan expansion to express a Cartesian tensor of rank n in terms of spherical tensors of varying ranks with the same kind of breakdown among the rank j tensors as we have found above for the spaces. An example of this is $\tau^{(2)} = \tau^{(1)} \otimes \tau^{(1)} = \mathbb{V}^{(2)} \oplus \mathbb{V}^{(1)} \oplus \mathbb{V}^{(0)}$; therefore, a rank 2 Cartesian tensor can be expressed as a sum of a spherical tensors of rank 2, 1, and 0. Explicitly, the formula will be

$$\bigotimes_{k=1}^{n} |j_k = 1, m_k\rangle = \sum_{j_p} \sum_{m=-j_p}^{j_p} |j_p, m\rangle\langle j_p, m| \left(\bigotimes_{k=1}^{n} |j_k = 1, m_k\rangle \right)$$

where the j_p that the first sum runs over is determined by n and the breakdown of $\tau^{(n)}$ into $\mathbb{V}^{(j)}$. There is a $_p$ index on j_p because there may be multiple orthogonal copies of the same $\mathbb{V}^{(j)}$ subspace in the direct sum; p indexes them. Of course, it will be tremendously tedious to calculate all these coefficents for $n > 2$, but it is feasible. There are tables of them, called the Wigner $6j$, $9j$, and $12j$ symbols, which provide the CG coefficients for addition of 3, 4, and 5 angular momenta respectively; clearly, these are applicable for Cartesian tensors of rank 3, 4, and 5.