Ph125:
Quantum Mechanics

Sunil Golwala

Fall 2007/Winter 2008

Contents

Contents

Contents

Contents

Contents

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

Revision Date: 2007/10/03

Section 1
Introduction to Course

Course Logistics

- ▶ The course webpage is

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

    All course logistics and assignments will be announced on this page. Check it regularly!

- ▶ 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!

- ▶ Problem sets: one per week, 4-6 problems. Posted via web site. Due date – M in class? Tu 10 am? Tu 5 pm? 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.

Course Logistics (cont.)

▶ 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 Sunday night 9-11 pm. There will be a TA office hour also.

▶ TAs: Hsin-Hua Lai, Chan Youn Park, Jaewon Song. Primarily responsible for solution sets and grading. Will rotate through TA office hour unless there is a strong preference for one.

Course Material

In this course, you will truly learn to attack basic quantum 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. We will also explore some of the interesting and unusual implications of quantum mechanics.

Course Material (cont.)

Fundamentals of QM:

- ▶ 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).
- ▶ Density matrix and thermal systems: how to mesh quantum states and classical thermal probability distributions.
- ▶ Symmetries: esp. how symmetries of the Hamiltonian determine conserved observables.
- ▶ Going into 3D: Rotations and coordinate angular momentum. How to use the angular momentum observables to classify 3D states.
- ▶ Formalism for spin angular momentum.

Course Material (cont.)

- ▶ 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.
- ▶ Time-dependendent perturbation theory: How to approach problems with small, time-dependent non-commuting terms; esp., electromagnetic radiation.
- ▶ Approximate methods for problems without exact solutions: variational methods, WKB approximation.
- ▶ Connections to classical mechanics: classical limits, symmetries, Hamiltonian formalism, Hamilton-Jacobi equation, WKB approximation, path integrals.

Course Material (cont.)

Applications:

- ▶ Hydrogen atom, including perturbations.
- ▶ Electromagnetic radiation of atoms.
- ▶ Scattering of particles.
- ▶ Second half of Ph125b is still open to requests. Some ideas:
    - ▶ Fermi gas
    - ▶ Phonons
    - ▶ Periodic potentials/tunneling resonances/Bloch waves/band structure
    - ▶ Bose-Einstein Condensation/Superconductivity
    - ▶ NMR/Rabi Oscillations/Two-Level Systems
    - ▶ Second quantization and quantum field theory
    - ▶ Squeezed states and quantum harmonic oscillators
    - ▶ Einstein-Podolsky-Rosen Experiment/Bell's inequalities
    - ▶ Quantum computing
    - ▶ Quantum cryptography
    - ▶ Quantum entanglement
    - ▶ Quantum measurement theory

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.

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.

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:

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

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 are playing fast and loose with δ functions here; we will make this rigorous 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:

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

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 $p = P |\psi\rangle$.

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

$$\Omega(X, P) = \omega(x \to X, p \to P) \qquad (2.4)$$

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.

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 $x \to X$ and $p \to P$.

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 cover path integrals and try to connect quantum mechanics to classical mechanics at a technical level.

Section 3
Mathematical Preliminaries

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 also explore the operator analogues of symmetry transformations in classical mechanics; while these do not correspond to physcial observables directly, we will see that they are *generated* by physical observables.
- ▶ 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.

Mathematical Preliminaries: Prologue

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 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:

▶ There is an addition operation on the vectors (vector addition) that is closed, meaning that addition of two vectors gives a third:

> 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?

▶ The vector addition operation is commutative and associative:

> $|v\rangle + |w\rangle = |w\rangle + |v\rangle$ and
> $(|v\rangle + |w\rangle) + |u\rangle = |v\rangle + (|w\rangle + |u\rangle)$

Linear Vector Spaces: Definitions (cont.)

▶ There is a unique zero or null vector $|0\rangle$ that, when added to a vector, gives back the vector:

> There exists a unique vector $|0\rangle$ in the vector space such that $|v\rangle + |0\rangle = |v\rangle$ for any vector $|v\rangle$.

Technically: there is a unique *identity* vector for the addition operation.

▶ Every vector has a unique inverse vector that gives the null vector when the vector and its inverse are added together:

> 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.)

▶ 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 (scalar addition):

$$(\alpha + \beta)|v\rangle = \alpha|v\rangle + \beta|v\rangle$$

▶ Scalar-vector multiplication is associative in the obvious way over multiplication in the field (scalar multiplication):

$$\alpha\left(\beta|v\rangle\right) = (\alpha\beta)|v\rangle$$

A few notes on what may or may not be inferred from the above:

▶ The addition rule in the field (scalar addition) must be consistent with the scalar-vector multiplication and vector addition rules; or, alternatively one may **define** scalar addition by the above distributivity and associativity rules. In either case, we draw the following conclusions without having to state them as axioms (proofs to be homework problems):

  ▶ Scalar addition is commutative and associative:

  $$[\alpha + (\beta + \gamma)] \, |v\rangle = [(\alpha + \beta) + \gamma] \, |v\rangle$$
  $$(\alpha + \beta) \, |v\rangle = (\beta + \alpha) \, |v\rangle$$

  ▶ The scalar addition identity 0 is consistent with the vector addition identity

  $$0|v\rangle = |0\rangle$$

  ▶ Scalar-vector multiplication against the vector addition identity yields the obvious result

  $$\alpha|0\rangle = |0\rangle$$

- ▶ The multiplication rule in the field (scalar multiplication) must be consistent with the scalar-vector multiplication rule; or, one may **define** scalar multiplication by the above scalar-vector multiplication associativity rule. Again, in either case, we can draw some conclusions without having to state them as axioms (and, again, the proofs will be homework):

  - ▶ Scalar multiplication is associative:

  $$[\alpha(\beta\gamma)]\,|v\rangle = [(\alpha\beta)\,\gamma]\,|v\rangle$$

  - ▶ Scalar multiplication need not have an identity or inverses.
  - ▶ Scalar multiplication need not be commutative.
  - ▶ But, if scalar multiplication does have an identity 1 and inverses, then they must again be consistent with scalar-vector multiplication and scalar addition, implying:

    - ▶ $1|v\rangle = |v\rangle$
    - ▶ $(-1)\,|v\rangle = -|v\rangle$
    - ▶ Scalar multiplication by 1 or -1 is commutative, but scalar multiplication between any other two scalars need not be.

Linear Vector Spaces: Definitions

Example 3.1: The set or real, antisymmetric $N \times N$ square matrices with the real numbers as the field.

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

▶ The sum of two antisymmetric matrices is pretty clearly antisymmetric.

▶ Addition of matrices is commutative and associative because the element-by-element addition operation is.

▶ The null vector is simply the matrix with all zeros.

▶ The additive inverse of a matrix is obtained by simply taking the additive inverse of each matrix 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 multiplied every element of the matrix one-by-one.

▶ Scalar-vector multiplication is associative for the same reason.

Lecture 2:
Vector Spaces continued
Inner Product Spaces

Revision Date: 2007/10/04

Linear Vector Spaces: Linear Independence and Bases

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 $\{|v_i\rangle\}$ is linearly independent if no one of them can be written in terms of the others. Mathematically, we state this as: there is no relation of the form

$$\sum_{i=1}^{n} \alpha_i |v_i\rangle = 0 \tag{3.1}$$

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

$$|v_1\rangle = \frac{1}{\alpha_1} \sum_{i=2}^{n} \alpha_i |v_i\rangle \tag{3.2}$$

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 that can be found has n members.

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 $\{|v_i\rangle\}$ in a n-dimensional vector space, any other vector $|v\rangle$ in the vector space can be expanded in terms of them:

$$|v\rangle = \sum_i \alpha_i |v_i\rangle \tag{3.3}$$

▶ 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 $\{\alpha_i\}$ for a particular vector $|v\rangle$ are called the components of $|v\rangle$. The vector space is said to be the space spanned by the basis.

Example 3.2: The set of real antisymmetric $N \times N$ matrices (with the real numbers as the field).

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

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

---

[1] $A^T = -A$ where ^{T} indicates matrix transposition.

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

Two notes:

- ▶ It is not necessary to assume $\langle v|v\rangle$ is real; it is implied by the transpose property.

- ▶ The above also implies antilinearity, $\langle \alpha v + \beta w|u\rangle = \alpha^*\langle v|u\rangle + \beta^*\langle w|u\rangle$

- ▶ The scalar result of the inner product need not be a member of the field of the vector space, though in most practical cases (and all cases we will consider) it is because this makes it algorithmically simple to calculate.

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 $\langle v|v\rangle^{1/2}$, which we write as $|v|$. A vector is normalized if its norm is 1; such a vector is termed a unit vector.

Two vectors are orthogonal or perpendicular if their inner product vanishes.

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.

Mathematical Preliminaries: Inner Product Spaces

Inner Product Spaces: Definitions (cont.)

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. Since all other vectors can be expanded in terms of the basis vectors, and the inner product is linear and antilinear, it is then straightforward to calculate the inner product of any two vectors. That is, if $\{|i\rangle\}$ is a basis (not necessarily orthonormal), and $|v\rangle = \sum_{i=1}^{n} v_i|i\rangle$ and $|w\rangle = \sum_{i=1}^{n} w_i|i\rangle$, then

$$\langle v\,|\,w\rangle = \sum_{ij} v_i^* w_j \langle i\,|\,j\rangle \tag{3.4}$$

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

$$\langle v\,|\,w\rangle = \sum_{ij} v_i^* w_j \delta_{ij} = \sum_{i} v_i^* w_i \tag{3.5}$$

and, for an inner product space defined such that component values can only be real numbers, such as 3D space, we just have the standard dot product.

Example 3.3: The set of real antisymmetric $N \times N$ matrices (with the real numbers as the field) with element-bpy-element multiplication and summing as the inner product.

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

$$\langle A | B \rangle = \sum_{kl} A_{kl} B_{kl} \tag{3.6}$$

where kl indicates the element in the kth row and lth column. Does it 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: Definitions (cont.)

Dirac Notation

Given a vector $|v\rangle$ with n components $\{v_i\}$ in terms of an orthonormal basis $\{|i\rangle\}$, we choose to make the following correspondences as **definitions**:

$$|v\rangle \longleftrightarrow \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \qquad \langle v| \longleftrightarrow \begin{bmatrix} v_1^* & v_2^* & \cdots & v_n^* \end{bmatrix} \qquad (3.7)$$

where the first quantity is called ket v (it is just a vector) and the second is called bra v. Separately, the space of all kets is a vector space and the space of all bras is a vector space. They do not intermingle when we only consider them as vector spaces because we have provided no rule for adding a bra and a ket.

You can check $|\alpha v\rangle = \alpha|v\rangle$ and $\langle\alpha v| = \langle v|\alpha^*$ by using the definitions. The placement of α^* to the right of $\langle v|$ is arbitrary and is done only for notational consistency with taking the adjoint (below).

Of course, we chose these definitions because the inner product calculation algorithm corresponds to simple matrix multiplication; *i.e.*, for $|v\rangle = \sum_{i=1}^{n} v_i |i\rangle$ and $|w\rangle = \sum_{i=1}^{n} w_i |i\rangle$, we have

$$\langle v | w \rangle = \sum_i v_i^* w_i = \begin{bmatrix} v_1^* & v_2^* & \cdots & v_n^* \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \langle v || w \rangle \qquad (3.8)$$

We note that the basis vectors satisfy

$$
|i\rangle \longleftrightarrow
\begin{bmatrix}
0 \\
\vdots \\
0 \\
1 \\
0 \\
\vdots \\
0
\end{bmatrix}
\qquad
\langle i | \longleftrightarrow
\begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix}
\tag{3.9}
$$

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

$$
|v\rangle = \sum_i v_i |i\rangle = \sum_i \langle i | v \rangle |i\rangle
\qquad
\langle v | = \sum_i \langle i | v_i^* = \sum_i \langle i | \langle v | i \rangle
\tag{3.10}
$$

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

Inner Product Spaces: Definitions (cont.)

Adjoints

We **define** the process of converting from bra to ket and vice versa as taking the adjoint. $|v\rangle$ is the adjoint of $\langle v|$ and vice versa. The simple algorithm for this is to 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 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.

Inner Product Spaces: Definitions (cont.)

▶ 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.) 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

$$|\vec{w}\,\rangle = \alpha_1 \cdots \alpha_l \langle w_1 \,|\, v \,\rangle \cdots \langle w_m \,|\, v_m \,\rangle |\, v \,\rangle \qquad (3.11)$$

where the $\{\alpha_i\}$, $\{\vec{w}_i\}$, and $\{|v_i\rangle\}$ are arbitrary (the index matchups mean nothing), we have that the adjoint is

$$\langle w \,| = \langle v \,|\langle v_1 \,|\, w_1 \,\rangle \cdots \langle v_m \,|\, w_m \,\rangle \alpha_1^* \cdots \alpha_l^* \qquad (3.12)$$

▶ vector expansions: We may write our vector expansions as

$$|v\,\rangle = \sum_i \langle i\,|\,v\,\rangle|i\,\rangle = \sum_i |i\,\rangle\langle i\,|\,v\,\rangle \qquad \langle v\,| = \sum_i \langle v\,|\,i\,\rangle\langle i\,| \qquad (3.13)$$

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.

Inner Product Spaces: Theorems

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"

$$|v + w|^2 = |v|^2 + |w|^2 + 2\,\mathcal{R}\left(\langle v\,|w\rangle\right) \tag{3.14}$$

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 $|w\rangle$ are 3-dimensional real vectors.

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. Draw it and it will become obvious.

Inner Product Spaces: Theorems (cont.)

Schwarz Inequality

$$|\langle v \,|\, w \rangle|^2 \leq |v|^2 |w|^2 \tag{3.15}$$

This is proven by applying the law of cosines to the vector

$$|v'\rangle = |v\rangle - \langle w \,|\, v \rangle \frac{|w\rangle}{|w|^2} = |v\rangle - \langle \widehat{w} \,|\, v \rangle |\widehat{w}\rangle \tag{3.16}$$

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

$$|v'\rangle = |v\rangle - |\widehat{w}\rangle\langle \widehat{w} \,|\, v \rangle \tag{3.17}$$

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.

Inner Product Spaces: Theorems (cont.)

$$|v + w| \leq |v| + |w| \tag{3.18}$$

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.

Lecture 3:
Inner Product Spaces continued
Linear Operators

Revision Date: 2007/10/05

Inner Product Spaces: Gram-Schmidt Orthogonalization

Gram-Schmidt Orthogonalization

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

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

We start out by taking the first new basis ket to be the first ket from the original set,

$$|1'\rangle = |v_1\rangle \tag{3.19}$$

Then, the second member of the orthogonal set is

$$|2'\rangle = |v_2\rangle - \frac{|1\rangle\langle 1|v_2\rangle}{\langle 1|1\rangle} \tag{3.20}$$

and so on; the generic formula is

$$|i'\rangle = |v_i\rangle - \sum_{j=1}^{i-1} \frac{|j\rangle\langle j|v_i\rangle}{\langle j|j\rangle} \tag{3.21}$$

Once one has an orthogonal set, it is easy to individually normalize each member, $|i\rangle = |i'\rangle / \langle i'|i'\rangle^{1/2}$.

One proves this theorem inductively, showing that if the first i-1 vectors have been orthogonalized, then the ith vector created via the above formula is orthogonal to the first i-1.

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

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

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

Example 3.4: Real antisymmetric, imaginary symmetric, and anti-Hermitian matrices.

The aforementioned set of real $N \times N$ matrices with a real number field form a subspace of the set of anti-Hermitian matrices[a], also with a real field, if we redefine the inner product to be

$$\langle A | B \rangle = \sum_{kl} A_{kl}^* B_{kl}$$

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.

---

[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 under a different inner product rule. That still holds for the above inner product rule because complex conjugation has no effect. With the new formula, purely imaginary symmetric matrices are also an inner product space because the complex conjugation 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.

Linear Operators

An operator Ω transforms a vector into another (possibly the same in some cases) vector, $\Omega|v\rangle = |w\rangle$.

An operator can also transform bras since they are just a different representation of the same vectors; $\langle u| = \langle v|\Omega$. But 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).

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

$$\Omega\left(\alpha|v\rangle + \beta|w\rangle\right) = \alpha\Omega|v\rangle + \beta\Omega|w\rangle \qquad \left(\langle v|\alpha + \langle w|\beta\right)\Omega = \alpha\langle v|\Omega + \beta\langle w|\Omega \tag{3.22}$$

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.

Linear Operators (cont.)

Example 3.5: Assorted operator examples

Shankar discusses the identity operator, which acts in any vector space, and the 3D rotation operator that acts on real 3D vectors. Another example is Gram-Schmidt orthogonalization, which maps a set of basis vectors in a set of normalized, orthogonal basis vectors. To be clear: there is not a single Gram-Schmidt operator, but clearly one can construct a particular Gram-Schmidt operator based on the initial basis set $|i'\rangle$; that operator maps the initial basis set into an orthonormal basis set, and does something else to the rest of the vector space.

Some operators we can derive from operators: products, commutators $[\Omega, \Lambda] = \Omega\Lambda - \Lambda\Omega$ (which are in general nonzero because most operators do not commute), and inverses (which are unique and commutative if they exist). See Shankar for details.

Linear Operators (cont.)

Linear Operator Action on Basis Vectors, Matrix Elements

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 $\{|i\rangle\}$ and a vector $|v\rangle = \sum_i v_i |i\rangle$ expanded in terms of them, we may write

$$\Omega|v\rangle = \Omega \sum_i v_i |i\rangle = \sum_i v_i \Omega |i\rangle \tag{3.23}$$

It is useful to break $\Omega|i\rangle$ into components by rewriting the expansion using $\sum_j |j\rangle\langle j|$:

$$\Omega|v\rangle = \sum_i v_i \Omega|i\rangle = \sum_i v_i \sum_j |j\rangle\langle j|\Omega|i\rangle \tag{3.24}$$

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

$$\Omega|v\rangle = \sum_{ij} \Omega_{ji} v_i |j\rangle \tag{3.25}$$

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

Linear Operators (cont.)

Of course, the above expression looks like matrix multiplication of a $n \times n$ matrix against a single-column n-row matrix:

$$[\Omega|v\rangle]_j = \langle j|\Omega|v\rangle = \sum_i \Omega_{ji} v_i \tag{3.26}$$

This makes sense: we were able to represent our vectors via single-column matrices (kets) or 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 ji element (jth row, ith column) is just the projection of the action of Ω on $|i\rangle$ onto $|j\rangle$.

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

$$\langle v|\Omega = \sum_i \langle i|\Omega v_i^* = \sum_{ij} v_i^* \langle i|\Omega|j\rangle\langle j| = \sum_{ij} v_i^* \Omega_{ij}\langle j| \tag{3.27}$$

$$\text{or} \quad [\langle v|\Omega]_j = \sum_i v_i^* \Omega_{ij} \tag{3.28}$$

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_{ij} is the projection of the action of Ω on $\langle i|$ onto $\langle j|$ (note the transposition of the indices relative to the ket case).

Linear Operators (cont.)

Projection Operators

We have already indirectly seen projection operators when we observed that vector expansions could be written in the form

$$|v\rangle = \sum_i v_i |i\rangle = \sum_i \langle i|v\rangle |i\rangle = \sum_i |i\rangle \langle i|v\rangle \tag{3.29}$$

The object $|i\rangle\langle i|$ acts on vectors, extracting the projection of $|v\rangle$ along $|i\rangle$ and then multiplying it by $|i\rangle$ to recover a vector rather than a scalar. This operator $P_i = |i\rangle\langle i|$ is called a projection operator because it projects out the piece of $|v\rangle$ along $|i\rangle$.

For a linear inner-product space and an orthonormal basis $\{|i\rangle\}$, it always holds that

$$\sum_i |i\rangle \langle i| = \sum_i P_i = I \tag{3.30}$$

where I is the identity operator. This simply states the fact that the sum of all the projections of a vector along an orthonormal basis is the vector itself. This of course need not work for non-orthonormal bases.

Linear Operators (cont.)

Projection operators have a very simply matrix representation:

$$(P_i)_{kl} = \langle k | P_i | l \rangle = \langle k | i \rangle \langle i | l \rangle = \delta_{ki} \delta_{il} \tag{3.31}$$

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

$$|v\rangle\langle w| = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \begin{bmatrix} w_1^* & w_2^* & \cdots & w_n^* \end{bmatrix} = \begin{bmatrix} v_1 w_1^* & v_1 w_2^* & \cdots & v_1 w_n^* \\ v_2 w_1^* & v_2 w_2^* & \cdots & v_2 w_n^* \\ \vdots & & \ddots & \vdots \\ v_n w_1^* & v_n w_2^* & \cdots & v_n w_n^* \end{bmatrix} \tag{3.32}$$

$$\text{or} \quad [|v\rangle\langle w|]_{kl} = v_k w_l^* \tag{3.33}$$

because, for a projection operator $P_i = |i\rangle\langle i|$, we have $v_k = \delta_{ki}$ and $w_l = \delta_{il}$ as shown earlier.

A projection operator need not project only along the vectors of an orthonormal basis. A projection operator can project along any direction, or even into any subspace of the space. The defining characteristic of a projection operator P is that $P^2 = P$: that applying the projection operator a second time results in no effect. This makes sense: once you have projected into a subspace, any further projection has no effect. It might perhaps be better to call the projection operators we defined above to be basis projection operators because they are a special case: they project along the basis vectors that define the matrix representation currently in use.

Note that $P^2 = P$ does not imply P is its own inverse: projection operators always have zero determinant and are hence noninvertible except when $P = I$.

Linear Operators (cont.)

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:

$$[\Omega\Lambda]_{ij} = \langle i|\Omega\Lambda|j\rangle = \langle i|\Omega\left(\sum_k |k\rangle\langle k|\right)\Lambda|j\rangle = \sum_k \langle i|\Omega|k\rangle\langle k|\Lambda|j\rangle = \sum_k \Omega_{ik}\Lambda_{kj}$$

(3.34)

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

Linear Operators (cont.)

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 $\langle v|$:

$$\text{if} \quad |w\rangle = \Omega|v\rangle \quad \text{then} \quad \langle w| = \langle v|\Omega^\dagger \tag{3.35}$$

defines $\Omega^\dagger$. To get an algorithmic formula for obtaining $\Omega^\dagger$ from Ω, we go to matrix representation:

$$\left[\Omega^\dagger\right]_{ij} = \langle i|\Omega^\dagger|j\rangle = (\Omega|i\rangle)^\dagger|j\rangle = \langle \Omega i|j\rangle = (\langle j|\Omega i\rangle)^* = (\langle j|\Omega|i\rangle)^* = \Omega_{ji}^* \tag{3.36}$$

So, one simply transposes and complex conjugates Ω's matrix representation to get the matrix representation of $\Omega^\dagger$.

Linear Operators (cont.)

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:

$$[\Omega\Lambda]^\dagger = \Lambda^\dagger\Omega^\dagger \tag{3.37}$$

which one easily proves by acting on an arbitrary vector:

$$(\left[\Omega\Lambda\right]|v\rangle)^\dagger = (\Omega\left[\Lambda|v\rangle\right])^\dagger = (\Lambda|v\rangle)^\dagger\,\Omega^\dagger = \left(\langle v|\Lambda^\dagger\right)\Omega^\dagger = \langle v|\Lambda^\dagger\Omega^\dagger \tag{3.38}$$

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.

Linear Operators (cont.)

Hermitian, Anti-Hermitian, and Unitary Operators: the Operators of QM

A Hermitian operator Ω has $\Omega^\dagger = \Omega$. An anti-Hermitian operator has $\Omega^\dagger = -\Omega$.

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.

You will have homework on products of Hermitian operators.

An unitary operator U has $U^\dagger = U^{-1}$.

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.

A product of unitary operators is unitary; one can see this by simply using the product rules for adjoints and inverses.

Linear Operators (cont.)

Inner Products and Unitary Operators

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

$$\text{if } |v'\rangle = U|v\rangle \text{ and } |w'\rangle = U|w\rangle \text{ then } \langle w'|v'\rangle = \langle w|v\rangle \tag{3.39}$$

The proof is trivial:

$$\langle w'|v'\rangle = (U|w\rangle)^{\dagger}(U|v\rangle) = \langle w|U^{\dagger}U|v\rangle = \langle w|v\rangle \tag{3.40}$$

We thus see that unitary operators are generalizations of rotation and other **orthogonal** operators from classical mechanics, which preserve the 3D 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:

$$\langle \text{column } i \,|\text{column } j\rangle = \sum_k U_{ki}^* U_{kj} = \sum_k \left[U^{\dagger}\right]_{ik} U_{kj} = \left[U^{\dagger}U\right]_{ij} = \delta_{ij} \tag{3.41}$$

The row version is similar.

Linear Operators (cont.)

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

$$\langle w'|\Omega|v'\rangle = (U|w\rangle)^{\dagger}\,\Omega\,(U|v\rangle) = \langle w|\left(U^{\dagger}\Omega U\right)|v\rangle \tag{3.42}$$

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

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.

Lecture 4:
Eigenvalue Problems

Revision Date: 2007/10/18

Eigenvalue Problems: Formalism

Motivation

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. We will therefore proceed quickly.

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

$$\Omega|v\rangle = \omega|v\rangle \quad \Longleftrightarrow \quad (\Omega - I\omega)|v\rangle = |0\rangle \tag{3.43}$$

(where I is the identity operator.) That is the statement of the problem.

Eigenvalue Problems: Formalism (cont.)

Solution of the Problem

Let's not belabor this. 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:

$$|\Omega - I\,\omega| = 0 \tag{3.44}$$

This equation is termed the characteristic equation for Ω.

To properly justify this, one must prove a general formula for the definition of the inverse of a matrix when a matrix representation is specified, which only is valid if the matrix determinant is nonzero, and then show that the above equation is inconsistent with the existence of an inverse, and hence the matrix determinant must vanish. This is done in Shankar Appendix A.1. In the spirit of this being a physics class, not a math class, we don't prove all of that, we just use the result.

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 (and hence the size of the matrices and vectors): 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

$$P^n(\omega) = \sum_{m=0}^{n} c_m \omega^m = 0 \tag{3.45}$$

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. As long as the field for the vector space is complex, then these are valid eigenvalues (if the field were real, we would simply 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.

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

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

$$(\Omega - I\omega_i)|\omega_i\rangle = |0\rangle \tag{3.46}$$

Since we explicitly know what the operator is – we know the elements of Ω and we know ω_i – all we need to do is solve for the elements of $|\omega_i\rangle$. 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 $|\omega_i\rangle$, leaving the overall normalization undetermined. The normalization of $|\omega_i\rangle$ is arbitrary since, if $|\omega_i\rangle$ is an eigenvector, then so will $\alpha|\omega_i\rangle$ for any α.

Even though we are assured that there are always n eigenvalues, it is not guaranteed that there are n eigenvectors. We can explicitly prove it when the matrix is Hermitian or unitary, as we will do below.

Example 3.6: In a 3D space with a complex field, rotation about the vector $\left(\widehat{x} + \widehat{y} + \widehat{z}\right)/\sqrt{3}$, which simply cyclically permutes the three unit vectors.

This matrix, which rotates $\widehat{x} \to \widehat{y}$, $\widehat{y} \to \widehat{z}$ and $\widehat{z} \to \widehat{x}$, is

$$A = \left[\begin{array}{ccc} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right] \tag{3.47}$$

The characteristic equation is

$$0 = |A - I\omega| = \left| \begin{array}{ccc} -\omega & 0 & 1 \\ 1 & -\omega & 0 \\ 0 & 1 & -\omega \end{array} \right| = -\omega^3 + 1 = 0 \tag{3.48}$$

$$\omega_1 = 1 \qquad \omega_2 = e^{2\pi i/3} \qquad \omega_3 = e^{-2\pi i/3} \tag{3.49}$$

Note that, if we had assumed a real field, we would have said that two of the eigenvalues did not exist.

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

$\omega_1 = 1$:

$$\left[\begin{array}{ccc} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{array} \right] \left[\begin{array}{c} v_{11} \\ v_{12} \\ v_{13} \end{array} \right] = |0\rangle \implies \begin{array}{c} -v_{11} + v_{13} = 0 \\ v_{11} - v_{12} = 0 \\ v_{12} - v_{13} = 0 \end{array} \tag{3.50}$$

$$\implies |\omega_1\rangle = \alpha \left[\begin{array}{c} 1 \\ 1 \\ 1 \end{array} \right] \tag{3.51}$$

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.

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

$$\begin{bmatrix} -e^{\frac{2\pi i}{3}} & 0 & 1 \\ 1 & -e^{\frac{2\pi i}{3}} & 0 \\ 0 & 1 & -e^{\frac{2\pi i}{3}} \end{bmatrix} \begin{bmatrix} v_{21} \\ v_{22} \\ v_{23} \end{bmatrix} = |0\rangle \implies \begin{array}{l} -e^{\frac{2\pi i}{3}} v_{21} + v_{23} = 0 \\ v_{21} - e^{\frac{2\pi i}{3}} v_{22} = 0 \\ v_{22} - e^{\frac{2\pi i}{3}} v_{23} = 0 \end{array} \quad (3.52)$$

$$\implies |\omega_2\rangle = \alpha \begin{bmatrix} 1 \\ e^{-\frac{2\pi i}{3}} \\ e^{\frac{2\pi i}{3}} \end{bmatrix} \quad (3.53)$$

Since all the elements are unit modulus, we again may take $\alpha = 1/\sqrt{3}$.

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

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

(3.54)

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

(3.55)

Again, we may take $\alpha = 1/\sqrt{3}$.

Eigenvalue Problems: Formalism (cont.)

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:

$$\text{if} \quad A|\omega, 1\rangle = \omega|\omega, 1\rangle \quad \text{and} \quad A|\omega, 2\rangle = \omega|\omega, 2\rangle \qquad (3.56)$$

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

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.

Eigenvalue Problems: Formalism (cont.)

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):

$$\langle\omega|\Omega|\omega\rangle = \omega\langle\omega|\omega\rangle \tag{3.57}$$

Also consider the adjoint of the above expression

$$\langle\omega|\Omega^\dagger|\omega\rangle = \omega^*\langle\omega|\omega\rangle \tag{3.58}$$

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

$$(\omega - \omega^*)\langle\omega|\omega\rangle = 0 \tag{3.59}$$

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

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

Given two eigenvalues ω_i and ω_j and corresponding eigenvectors $|\omega_i\rangle$ and $|\omega_j\rangle$, we have

$$\langle\omega_i|\Omega|\omega_j\rangle = \langle\omega_i|\omega_j|\omega_j\rangle = \omega_j\langle\omega_i|\omega_j\rangle \tag{3.60}$$

and

$$\langle\omega_i|\Omega|\omega_j\rangle = \left(\Omega^\dagger|\omega_i\rangle\right)^\dagger|\omega_j\rangle = (\Omega|\omega_i\rangle)^\dagger|\omega_j\rangle = (\omega_i|\omega_i\rangle)^\dagger|\omega_j\rangle = \langle\omega_i|\omega_i^*|\omega_j\rangle \tag{3.61}$$
$$= \langle\omega_i|\omega_i|\omega_j\rangle = \omega_i\langle\omega_i|\omega_j\rangle$$

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

$$\left(\omega_i - \omega_j\right)\langle\omega_i|\omega_j\rangle = 0 \tag{3.62}$$

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

Eigenvalue Problems: 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 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 roots. There are thus n eigenvalues, nondegenerate by assumption here.

We have show 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.

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 slightly different tack from Shankar. Clearly, if one writes Ω in terms of a matrix representation in which its eigenvectors are the basis, then its eigenvectors have matrix representation

$$
|1\rangle = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad |2\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \cdots \quad |n\rangle = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{3.63}
$$

Clearly, then, in order to have the appropriate action on its eigenvectors in this representation, Ω must have representation

$$
\Omega = \begin{bmatrix} \omega_1 & 0 & \cdots & 0 \\ 0 & \omega_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \omega_n \end{bmatrix} \tag{3.64}
$$

as expected.

Eigenvalue Problems: 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.

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.

Eigenvalue Problems: 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:

$$\langle\omega\,|\,\omega\rangle = \langle\omega\,|\,U^{\dagger}U|\omega\rangle = \langle\omega\,|\,\omega^{*}\omega|\omega\rangle \Longrightarrow (\omega^{*}\omega - 1)\,\langle\omega\,|\,\omega\rangle = 0 \qquad (3.65)$$

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 $|\omega_i\rangle$ and $|\omega_j\rangle$ of two nondegenerate eigenvalues $\omega_i \neq \omega_j$:

$$\langle\omega_i\,|\,\omega_j\rangle = \langle\omega_i\,|\,U^{\dagger}U|\omega_j\rangle = \langle\omega\,|\,\omega_i^{*}\omega_j|\omega\rangle \Longrightarrow (\omega_i^{*}\omega_j - 1)\,\langle\omega_i\,|\,\omega_j\rangle = 0 \qquad (3.66)$$

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

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

Eigenvalue Problems: Formalism (cont.)

Diagonalization of Hermitian Matrices

Since we have shown that one can always construct an orthonormal basis of the eigenvectors of a Hermitian matrix, we are also provided with an explicit matrix representation of a unitary operator that will transform from the original basis for the vector space to the basis in which Ω is diagonal; that matrix is:

$$U_\Omega^\dagger = \begin{bmatrix} \langle \omega_1 | \\ \langle \omega_2 | \\ \vdots \\ \langle \omega_n | \end{bmatrix} \tag{3.67}$$

where the notation means simply that each row of $U_\Omega^\dagger$ is the bra corresponding to the associated eigenvector of Ω. This form guarantees that $U_\Omega^\dagger$ will, for example, rotate $|\omega_1\rangle$ to become a column matrix with 1 in its first element and 0 otherwise; *i.e.*, the first orthonormalized basis vector defining the representation. We showed earlier that the unitarity condition is the equivalent of the rows or columns being orthonormal; we are guaranteed this will hold for $U_\Omega^\dagger$ because Ω is Hermitian and hence the eigenvectors are orthonormal.

Note that this definition is identical to Shankar's Equation 1.8.2. We use the definition

$$|i\rangle = U_\Omega^\dagger |\omega_i\rangle \tag{3.68}$$

while Shankar uses

$$|\omega_i\rangle = U_\Omega |i\rangle \tag{3.69}$$

How does U_Ω act on operators? Consider the following:

$$\langle\omega_j|\Omega|\omega_i\rangle = \langle\omega_j|U_\Omega\, U_\Omega^\dagger\, \Omega\, U_\Omega U_\Omega^\dagger|\omega_i\rangle = \langle j|U_\Omega^\dagger\, \Omega\, U_\Omega|i\rangle \tag{3.70}$$

Since we know $\langle\omega_j|\Omega|\omega_i\rangle = \omega_i\delta_{ij}$, it must therefore hold that the unitary transformation $U_\Omega^\dagger\, \Omega\, U_\Omega$ must render Ω diagonal.

A final note, though, on the difference between "arithmetic operations" and "transformations." The above notation really implies that $\Omega' = U_\Omega^\dagger \, \Omega \, U_\Omega$ is a **different operator** than Ω. But we have explained how whether Ω is diagonalized is merely a matter of representation – which basis set we choose to write it down in terms of. The operator Ω itself does not change because of choice of basis; **its representation changes**.

How can these two viewpoints be consistent? They are consistent because, in creating Ω', we have also changed its eigenvectors by $|i\rangle = U_\Omega^\dagger |\omega_i\rangle$. That is, the unitary operator changes both the operator and the eigenvectors is such a way that the new vectors $|i\rangle$ are the eigenvectors of the new operator Ω'.

In practice, the distinction between, on the one hand, transforming the operator and its eigenvectors, and, on the other hand, writing the operator and its eigenvectors in a representation in which the eigenvectors are diagonal, is lost. In practice, if we have the representation of Ω and its eigenvectors in a basis in which Ω is not diagonal, and we want their representations in the basis of Ω's eigenvectors, we apply the above unitary transformation **as purely an arithmetic procedure** to the matrix representation of the eigenvectors and the operator in the original basis to get the representation in the eigenbasis. (You can check that this correctly does the arithmetic needed to get the components in the eigenbasis from the components in the original basis.) The intent is not to change the operator and eigenvectors.

Lecture 5:
Eigenvalue Problems continued
Functions of Operators

Revision Date: 2007/10/12

Eigenvalue Problems: Formalism

If Ω and Λ are two commuting Hermitian operators, there is a basis of common eigenvectors that diagonalizes both.

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, $\{\omega_i\}$ and $\{|\omega_i\rangle\}$. Then we have

$$\Lambda\left[\Omega|\omega_i\rangle\right] = \omega_i\Lambda|\omega_i\rangle \tag{3.71}$$

Using the fact that Ω and Λ commute, we have

$$\Omega\left[\Lambda|\omega_i\rangle\right] = \omega_i\Lambda|\omega_i\rangle \tag{3.72}$$

So, if $|\omega_i\rangle$ is an eigenvector of Ω, so is $\Lambda|\omega_i\rangle$. Assuming no degeneracies, then $\Lambda|\omega_i\rangle$ must be just a multiple of $|\omega_i\rangle$, $\Lambda|\omega_i\rangle = \lambda_i|\omega_i\rangle$, because the eigenvector $|\omega_i\rangle$ is specified completely up to a multiplicative constant when the eigenvalues of Ω are nondegenerate. But the statement $\Lambda|\omega_i\rangle = \lambda_i|\omega_i\rangle$ says that $|\omega_i\rangle$ is an eigenvector of Λ, too, with eigenvalue λ_i. Hence, the eigenvectors of Ω, which are orthonormal and diagonalize Ω, are also a set of eigenvectors for Λ and thus which must diagonalize it, too.

Degenerate case:

Of course, as usual, one must think more in the degenerate case because, if ω_i is a degenerate eigenvalue, then the fact that $|\omega_i\rangle$ and $\Lambda|\omega_i\rangle$ are both eigenvectors of Ω with eigenvalue $|\omega_i\rangle$ does not imply $\Lambda|\omega_i\rangle = \lambda_i|\omega_i\rangle$; Λ could map $|\omega_i\rangle$ somewhere else in the degenerate subspace of eigenvectors of Ω with eigenvalue ω_i.

It is straightforward to deal with this. Let us consider three cases.

▶ **Λ has degenerate eigenvalues but Ω does not.**
Our proof carries through in this case because it is Ω's degeneracy that matters in the proof. Ω's eigenvectors provide a perfectly good set of eigenvectors for Λ, also, so Λ and Ω are diagonalized when Ω's eigenvectors are used as the basis.

▶ **Ω has degenerate eigenvalues but Λ does not.**
Simply exchange their roles – use Λ's nondegenerate eigenvectors as the diagonalizing basis.

- ▶ **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.

Eigenvalue Problems: A Worked Example

Example 3.7: 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 simply call out a few high points and refer you to the text for all the algebra.

The problem consists of two masses coupled to each other by a spring and to fixed walls on either end.
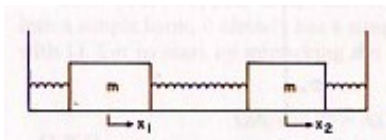
One of course writes Newton's Second Law for the system as a matrix equation, which is

$$\left[\begin{array}{c} \ddot{x}_1 \\ \ddot{x}_2 \end{array} \right] + \frac{k}{m} \left[\begin{array}{cc} 2 & -1 \\ -1 & 2 \end{array} \right] \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right] = \left[\begin{array}{c} 0 \\ 0 \end{array} \right] \tag{3.73}$$

One assumes a harmonic solution with time dependence $e^{i\omega t}$ (of which we will take the real part in the end), which lets one evaluate the time derivatives, leaving a matrix equation

$$\left(\left[\begin{array}{cc} 2 & -1 \\ -1 & 2 \end{array} \right] - \lambda \right) \left[\begin{array}{c} x_{1,0} \\ x_{2,0} \end{array} \right] e^{i\omega t} = \left[\begin{array}{c} 0 \\ 0 \end{array} \right] \tag{3.74}$$

where $\lambda = \omega^2/(k/m)$. We shall denote the above matrix operator by Λ. The equation is of course a characteristic equation of the form defined earlier in this section.

One proceeds with the solution as we outlined. One finds eigenvalues and eigenvectors:

$$\lambda_1 = 1 \qquad \omega_1 = \sqrt{\frac{k}{m}} \qquad |1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad (3.75)$$

$$\lambda_2 = 3 \qquad \omega_2 = \sqrt{3\frac{k}{m}} \qquad |2\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad (3.76)$$

An arbitrary state can be decomposed in terms of these eigenvectors:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{\sqrt{2}} (x_1 + x_2) |1\rangle + \frac{1}{\sqrt{2}} (x_1 - x_2) |2\rangle \qquad (3.77)$$

Eigenvalue Problems: A Worked Example (cont.)

The unitary matrix needed to rotate the matrix representation components from the coordinate basis to the eigenbasis is just the matrix whose rows are the eigenbras:

$$R^\dagger = \frac{1}{\sqrt{2}} \left[\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array} \right] \tag{3.78}$$

We can check this by applying it to an arbitrary state in the coordinate basis:

$$R^\dagger \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right] = \frac{1}{\sqrt{2}} \left[\begin{array}{c} x_1 + x_2 \\ x_1 - x_2 \end{array} \right] \tag{3.79}$$

One can check that the above rotates the eigenvector ($x_1 = 1/\sqrt{2}$, $x_2 = 1/\sqrt{2}$) to become the unit vector $(1, 0)$ as we expect from Equation 3.68. We apply the corresponding transformation to the operator Λ whose eigenvalues and eigenvectors we found:

$$R^\dagger \Lambda R = \frac{1}{\sqrt{2}} \left[\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array} \right] \left[\begin{array}{cc} 2 & -1 \\ -1 & 2 \end{array} \right] \frac{1}{\sqrt{2}} \left[\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array} \right] = \left[\begin{array}{cc} 1 & 0 \\ 0 & 3 \end{array} \right] \tag{3.80}$$

As expected, R diagonalizes our original operator Λ and the diagonal elements are its eigenvalues, $\lambda_1 = 1$ and $\lambda_2 = 3$.

Now, let's construct the full time evolution.

We will write an arbitrary vector describing the system at time t as $|v(t)\rangle$. If we decompose in terms of the eigenbasis, this state is written as

$$|v(t)\rangle = \left[\begin{array}{c} v_1(t) \\ v_2(t) \end{array} \right] = v_1(t)|1\rangle + v_2(t)|2\rangle \tag{3.81}$$

Note that the vectors $|1\rangle$ and $|2\rangle$ have no time dependence themselves because they were just solutions to the characteristic equation; the time dependence had been removed at that point. The time dependence will be carried by the coefficients in the expansion of $|v\rangle$ in terms of the eigenbasis.

How do $v_1(t)$ and $v_2(t)$ evolve? We have assumed a harmonic solution of the problem and diagonalized the matrix Λ with its eigenbasis. In that basis, the two harmonic solutions separate, with $e^{i\omega_1 t}$ attached to the $|1\rangle$ solution and $e^{i\omega_2 t}$ attached to the $|2\rangle$ solution. We will write these in real (cos and sin) form. Thus, we have that the arbitrary solution is

$$
\begin{aligned}
|v(t)\rangle = {} & \left(v_1(t=0)\cos(\omega_1 t) + \frac{1}{\omega_1}\dot{v}_1(t=0)\sin(\omega_1 t) \right) |1\rangle \qquad (3.82) \\
& + \left(v_2(t=0)\cos(\omega_2 t) + \frac{1}{\omega_2}\dot{v}_2(t=0)\sin(\omega_2 t) \right) |2\rangle
\end{aligned}
$$

In the above, the coefficient of each basis vector is just the standard general sinusoidal solution to the harmonic oscillator differential equation: cos is multiplied by the initial position and sin is multiplied by the initial velocity (and $1/\omega$).

For simplicity, we assume as Shankar does that the initial velocities vanish, so the above reduces to

$$|v(t)\rangle = v_1(t=0)\cos(\omega_1 t)|1\rangle + v_2(t=0)\cos(\omega_2 t)|2\rangle \qquad (3.83)$$

$$= \cos(\omega_1 t)|1\rangle\langle 1\,|v(t=0)\rangle + \cos(\omega_2 t)|2\rangle\langle 2\,|v(t=0)\rangle \qquad (3.84)$$

We are left with this simple form arises because the Λ operator is diagonal in this basis: **the eigenvectors time-evolve independently.** The eigenvectors are termed **normal modes** because, when the system is started in a normal mode, it is a simple sinusoidal oscillator at the frequency corresponding to that normal mode. These are the "natural oscillation modes" of the system. For the system we are studying, those modes correspond to both masses moving together (symmetrically) so the external springs expand and compress but the central one stays fixed; and to the masses moving by equal amounts in opposite directions (asymmetrically). Note that the symmetric mode is lower frequency; we shall see the same kind of behavior in QM: the symmetric mode will tend to be the lower energy one.

Though normal modes are convenient, we are interested in what x_1 and x_2 are, though, and we also specify initial conditions through $x_1(t=0)$ and $x_2(t=0)$, so we make use of the unitary matrix R twice to rewrite our expression in terms of the coordinates rather than the normal modes:

$$|v(t=0)\,\rangle = R^\dagger \left[\begin{array}{c} x_1(t=0) \\ x_2(t=0) \end{array} \right] \quad \text{and} \quad \left[\begin{array}{c} x_1(t) \\ x_2(t) \end{array} \right] = R|v(t)\,\rangle \qquad (3.85)$$

We thus rewrite our time evolution as

$$\left[\begin{array}{c} x_1(t) \\ x_2(t) \end{array} \right] = R\left[\cos(\omega_1 t)\,|1\,\rangle\langle 1\,| + \cos(\omega_2 t)\,|2\,\rangle\langle 2\,| \right] R^\dagger \left[\begin{array}{c} x_1(t=0) \\ x_2(t=0) \end{array} \right] \qquad (3.86)$$

Eigenvalue Problems: A Worked Example (cont.)

Now, you can find the above equation written out completely as Shankar Equation 1.8.39. It is not particularly elegant. What is elegant is that we can write the above very simply as

$$
\left[\begin{array}{c} x_1(t) \\ x_2(t) \end{array} \right] = U(t) \left[\begin{array}{c} x_1(t=0) \\ x_2(t=0) \end{array} \right] \tag{3.87}
$$

$$
U(t) = R \left[\cos(\omega_1 t)|1\rangle\langle 1| + \cos(\omega_2 t)|2\rangle\langle 2| \right] R^\dagger \tag{3.88}
$$

That, is once the initial state is specified in terms of our desired basis, the time evolution in that basis is specific by a single operator $U(t)$ that rotates from the coordinate basis to the eigenvector basis, does the simple time evolution (which is independent for each eigenvector) in this basis, then rotates back to the coordinate basis because that is where we want the final numbers.

The application to quantum mechanics is quite clear. The Schrödinger equation specifies time evolution in terms of the action of the Hamiltonian H on the state:

$$i\,\hbar\,\frac{d}{dt}|\psi\rangle = H|\psi\rangle \tag{3.89}$$

The time evolution of the eigenvectors of H will be very simple and will be specified by the eigenvalues of H. So, we will solve the eigenvalue problem of H, find the eigenvectors, and find the unitary matrix that rotates from the basis in which we would like to write the particle state to the eigenbasis. Then the time evolution will simply be

$$|\psi(t)\rangle = U(t)|\psi(t=0)\rangle \qquad U(t) = R\left(\sum_i u_i(t)|i\rangle\langle i|\right) R^\dagger \tag{3.90}$$

where $u_i(t)$ gives the time evolution of the ith eigenvector. $U(t)$ is known as the propagator because it propagates the initial conditions forward to an arbitrary point in time.

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

$$f(x) = \sum_{k=0}^{\infty} c_k x^k \tag{3.91}$$

then we simply make the obvious definition

$$f(\Omega) = \sum_{k=0}^{\infty} c_k \Omega^k \tag{3.92}$$

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,

$$\left[\Omega^2\right]_{ij} = \sum_k \Omega_{ik}\Omega_{kj} = \sum_k \omega_i\delta_{ik}\omega_j\delta_{kj} = \omega_i\omega_j\delta_{ij} = \omega_i^2\delta_{ij} \tag{3.93}$$

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

$$
e^{\Omega} = \sum_{k=0}^{\infty} \frac{1}{k!} \Omega^k = \sum_{k=0}^{\infty} \frac{1}{k!} \begin{bmatrix} \omega_1 & 0 & \cdots & 0 \\ 0 & \omega_2 & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & \omega_n \end{bmatrix}^k \tag{3.94}
$$

$$
= \begin{bmatrix} \sum_{k=0}^{\infty} \frac{1}{k!} \omega_1^k & 0 & \cdots & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{1}{k!} \omega_2^k & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & \sum_{k=0}^{\infty} \frac{1}{k!} \omega_n^k \end{bmatrix} \tag{3.95}
$$

$$
= \begin{bmatrix} e^{\omega_1} & 0 & \cdots & 0 \\ 0 & e^{\omega_2} & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & e^{\omega_n} \end{bmatrix} \tag{3.96}
$$

Related examples are sines and cosines and their hyperbolic counterparts.

Functions of Operators (cont.)

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:

$$e^{\alpha\Omega+\beta\Lambda} = \sum_{k=0}^{\infty} \frac{1}{k!}(\alpha\Omega+\beta\Lambda)^k \tag{3.97}$$

$$= I + (\alpha\Omega+\beta\Lambda) + \frac{1}{2}(\alpha\Omega+\beta\Lambda)^2 + \sum_{k=3}^{\infty} \frac{1}{k!}(\alpha\Omega+\beta\Lambda)^k \tag{3.98}$$

$$= I + (\alpha\Omega+\beta\Lambda) + \frac{1}{2}\left[\alpha^2\Omega^2 + \alpha\beta(\Omega\Lambda+\Lambda\Omega) + \beta^2\Lambda^2\right] + \sum_{k=3}^{\infty} \frac{1}{k!}(\alpha\Omega+\beta\Lambda)^k \tag{3.99}$$

$$e^{\alpha\Omega}e^{\beta\Lambda} = \left[\sum_{k=0}^{\infty} \frac{1}{k!}(\alpha\Omega)^k\right]\left[\sum_{m=0}^{\infty} \frac{1}{m!}(\beta\Lambda)^m\right] \tag{3.100}$$

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[2]. 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.

---

[2] 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.

Lecture 6:
Calculus with Operators
Infinite-Dimensional Generalization

Revision Date: 2007/10/12

Calculus with Operators

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:

$$\left[\frac{d}{d\lambda}e^{\lambda\Omega}\right]_{ij} = \frac{d}{d\lambda}\left[e^{\lambda\Omega}\right]_{ij} = \frac{d}{d\lambda}e^{\lambda\omega_i}\delta_{ij} = \omega_i e^{\lambda\omega_i}\delta_{ij} = \left[\Omega e^{\lambda\Omega}\right]_{ij} \qquad (3.103)$$

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

$$\frac{d}{d\lambda}e^{\lambda\Omega} = \Omega e^{\lambda\Omega} \qquad (3.104)$$

Of course, we could have placed Ω on the right side too since Ω and $e^{\lambda\Omega}$ commute.

▶ By power series:

$$\frac{d}{d\lambda}e^{\lambda\Omega} = \sum_{k=0}^{\infty} \frac{d}{d\lambda}\left[\frac{1}{k!}\lambda^k\Omega^k\right] = \sum_{k=1}^{\infty} \frac{1}{(k-1)!}\lambda^{k-1}\Omega^{k-1} \tag{3.105}$$

$$= \Omega \sum_{m=0}^{\infty} \frac{1}{m!}\lambda^m\Omega^m = \Omega e^{\lambda\Omega} \tag{3.106}$$

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.

Calculus with Operators (cont.)

And integration?

Integration is also a linear operation, so, it can always be written as element-by-element integration as we did with differentiation:

$$\left[\int_{\lambda_0}^{\lambda} d\lambda' \, \Theta(\lambda')\right]_{ij} = \int_{\lambda_0}^{\lambda} d\lambda' \left[\Theta(\lambda')\right]_{ij} = \int_{\lambda_0}^{\lambda} d\lambda' \, \Theta_{ij}(\lambda') \tag{3.107}$$

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:

$$\int_{\lambda_0}^{\lambda} d\lambda' \, \Omega \, e^{\lambda'\Omega} = \sum_{k=0}^{\infty} \int_{\lambda_0}^{\lambda} d\lambda' \, \Omega \left[\frac{1}{k!}\left(\lambda'\right)^k \Omega^k\right] = \sum_{k=0}^{\infty} \frac{1}{(k+1)!}\left(\lambda^{k+1} - \lambda_0^{k+1}\right)\Omega^{k+1} \tag{3.108}$$

$$= \sum_{m=0}^{\infty} \frac{1}{m!}\left(\lambda^m - \lambda_0^m\right)\Omega^m = e^{\lambda\Omega} - e^{\lambda_0\Omega} \tag{3.109}$$

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

$$\int_{\lambda_0}^{\lambda} d\lambda' \, e^{\lambda'\Omega} = \Omega^{-1}\left(e^{\lambda\Omega} - e^{\lambda_0\Omega}\right) \tag{3.110}$$

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.

Infinite-Dimensional Generalization: Examples

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:

  $$|n\rangle \longleftrightarrow x^n \tag{3.111}$$

  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.

Infinite-Dimensional Generalization: Examples (cont.)

▶ **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

From Finite to Infinite-Dimensional Vector Spaces

We will extend our analysis by considering functions on a set of discrete points on an interval on the real line, and then taking the limit as the number of such points is made infinite and the spacing made to vanish.

Consider the interval $[0, L]$. Let there be n points spaced out by $L/(n+1)$, $x_i = i\, L/n$ with $i = 1, \ldots, n$. Let our vector space be the set of functions on this discrete set of points, $\{f(\{x_i\})\}$. Clearly, we can represent such a function by a vector

$$|f\rangle = \left[\begin{array}{c} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_n) \end{array} \right] \tag{3.112}$$

It is fairly obvious that, with appropriate choice of the functions being real or complex and with real or complex coefficients, the set is closed and satisfies the usually vector space arithmetic rules.

An obvious basis for the space is the set of n functions that take on value 1 at one point and zero elsewhere:

$$|1\rangle = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad |2\rangle = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad \cdots \quad |n\rangle = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ n \end{bmatrix} \tag{3.113}$$

Then, any function f is simply written in ket form as

$$|f\rangle = \sum_{i=1}^{n} f(x_i)|i\rangle \tag{3.114}$$

The inner product is just the obvious matrix multiplication, which gives

$$\langle i|j\rangle = \delta_{ij} \tag{3.115}$$

From this, we infer that the inner product of two arbitary kets is

$$\langle f \,|\, g \rangle = \sum_{ij} f(x_i)\, g(x_j) \langle i \,|\, j \rangle = \sum_{i} f(x_i)\, g(x_i) \tag{3.116}$$

and the norm is thus

$$\langle f \,|\, f \rangle = \sum_{i} [f(x_i)]^2 \tag{3.117}$$

The standard outer product rule gives us that

$$\sum_{i=1}^{n} |i\rangle\langle i| = I \tag{3.118}$$

the identity operator. This is also knows as a completeness relation. It will prove important below.

Lecture 7:
Infinite-Dimensional Generalization continued

Revision Date: 2007/10/15

Infinite-Dimensional Generalization: From Finite to Infinite Dimensions

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:

$$\langle f \,|\, g \rangle = \sum_{i=1}^{n} f(x_i)\, g(x_i)\Delta \quad \text{with} \quad \Delta = \frac{L}{n+1} \tag{3.119}$$

Now, as we let $n \to \infty$, we recognize that the sum converts to an integral:

$$\lim_{n \to \infty} \sum_{i=1}^{n} f(x_i) g(x_i)\Delta = \int_0^L dx\, f(x)\, g(x) \tag{3.120}$$

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 i anymore, we must now label the points on which the function is defined by their position x. For complex functions and an arbitrary interval $[a, b]$ we may generalize this to

$$\langle f \,|\, g \rangle = \int_a^b dx\, f^*(x)\, g(x) \tag{3.121}$$

With the transition from the index i to position x, we also must sort out what happens to our basis kets. Let's 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 still hold. It now takes the form

$$\int_a^b dx' \, |x'\rangle\langle x'| = I \tag{3.122}$$

(The change of dummy variable from x to x' facilitates the next step.) Let's apply this to $\langle x |$ on the left an arbitrary ket $|f\rangle$ on the right:

$$\int_a^b dx' \langle x | x' \rangle\langle x' | f \rangle = \langle x | I | f \rangle = \langle x | f \rangle \tag{3.123}$$

We would certainly like to still have $\langle x | f \rangle = f(x)$ as we had in the finite case. Requiring this turns the above equation into a condition on $\langle x | x' \rangle$:

$$\int_a^b dx' \langle x | x' \rangle f(x') = f(x) \tag{3.124}$$

Infinite-Dimensional Generalization: From Finite to Infinite Dimensions (cont.)

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 $a < x < b$):

$$f(x) = \int_a^b dx' \langle x | x' \rangle f(x') = \lim_{\epsilon \to 0} \int_{x-\epsilon}^{x+\epsilon} dx' \langle x | x' \rangle f(x') \qquad (3.125)$$

$$= f(x) \lim_{\epsilon \to 0} \int_{x-\epsilon}^{x+\epsilon} dx' \langle x | x' \rangle \qquad (3.126)$$

where the last step is possible assuming $f(x)$ is continuous at x. Since $f(x)$ was arbitrary, we therefore have a generic requirement on $\langle x | x' \rangle$:

$$1 = \lim_{\epsilon \to 0} \int_{x-\epsilon}^{x+\epsilon} dx' \langle x | x' \rangle \quad \text{for} \quad a < x < b \qquad (3.127)$$

which needs to be coupled with our orthogonality requirement

$$\langle x | x' \rangle = 0 \quad \text{for} \quad x \neq x' \qquad (3.128)$$

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.

Let's briefly summarize our continuous x limit:

▶ Our space consists of all the functions on the interval $[a, b]$. $|f\rangle$ designates the vector space member **corresponding to** the function $f(x)$. The significance of the boldface is that the relation between $|f\rangle$ and $f(x)$ is not one of equality, but of one-to-one correspondence.

▶ We define the inner product between two members of the space $f(x)$ and $g(x)$ as

$$\langle f | g \rangle = \int_a^b dx\, f^*(x)\, g(x) \tag{3.129}$$

▶ We take as an orthonormal basis the kets $|x\rangle$ that are defined by the requirement

$$\langle x | x' \rangle = \delta(x - x') \quad \Longleftrightarrow \quad \begin{array}{l} \langle x | x' \rangle = 0 \quad \text{for} \quad x \neq x' \\ \int_{x-\epsilon}^{x+\epsilon} dx'\, \langle x | x' \rangle = 1 \end{array} \tag{3.130}$$

▶ (This one is implied by the above, it is not independent.)
The function $f(x)$ is returned by taking the inner product of $|f\rangle$ with the bra $\langle x|$:

$$f(x) = \langle x|f\rangle \tag{3.131}$$

This is an important point. $|f\rangle$ 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 $|x\rangle$. 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 $f(x)$. This point is why we have not somehow given an explicit expression for $|x\rangle$: there is no explicit expression for $|f\rangle$ either!

▶ (Again, this is implied by the above.)
We have the two facts:

$$\langle x | x' \rangle = \delta(x - x') \qquad f(x) = \langle x | f \rangle \tag{3.132}$$

Therefore, it is reasonable to make the correspondence

$$|x'\rangle \longleftrightarrow \delta(x - x') \tag{3.133}$$

just as we make the correspondence

$$|f\rangle \longleftrightarrow f(x) \tag{3.134}$$

Thus, we have defined $|x'\rangle$ as explicitly as we have defined $|f\rangle$.

This $\{|x\rangle\}$ basis and the distinction between $|f\rangle$ and $f(x)$ is usually very difficult to get one's mind around the first time because you are not used to thinking that the function f has any meaning besides $f(x)$. We are now saying it does.

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 3.127 nonzero.

You should be offended by such an object; to improve your acceptance of this function, think of it as a limit of a reasonable function:

$$\delta(x) = \lim_{\Delta \to 0} \begin{cases} \frac{1}{\Delta} & |x| < \frac{\Delta}{2} \\ 0 & |x| \geq \frac{\Delta}{2} \end{cases} \tag{3.135}$$

$$\delta(x) = \lim_{\Delta \to 0} \frac{1}{\sqrt{\pi \, \Delta^2}} \exp\left(-\frac{x^2}{\Delta^2}\right) \tag{3.136}$$

Infinite-Dimensional Generalization: Properties of the Dirac Delta Function (cont.)

We will later find it useful to discuss derivatives of the delta function, so let's work out what they are. *For those who worry about the validity of the following manipulations, you may repeat them on the aforementioned approximations to the delta function along with the appropriate limiting procedure.* The first derivative can be obtained via integration by parts:

$$\int_{x-\epsilon}^{x+\epsilon} dx' \left[\frac{d}{dx} \delta(x-x') \right] f(x') = \frac{d}{dx} \int_{x-\epsilon}^{x+\epsilon} dx' \delta(x-x') f(x') = \frac{d}{dx} f(x) \quad (3.137)$$

(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 may write this concisely in the following form:

$$\frac{d}{dx} \delta(x-x') = \delta(x-x') \frac{d}{dx'} \quad (3.138)$$

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'.

Infinite-Dimensional Generalization: Properties of the Dirac Delta Function (cont.)

What about $\frac{d}{dx'}\delta(x - x')$? We can figure this out by integration by parts:

$$\int_{x-\epsilon}^{x+\epsilon} dx' \left[\frac{d}{dx'} \delta(x - x') \right] f(x') = \left[\delta(x - x')\, f(x') \right]\big|_{x-\epsilon}^{x+\epsilon} \tag{3.139}$$
$$- \int_{x-\epsilon}^{x+\epsilon} dx' \delta(x - x')\, \frac{d}{dx'}\, f(x')$$
$$= -\frac{d}{dx}\, f(x) \tag{3.140}$$

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 value at the two endpoints becomes equal and infinite and so the term vanishes) and we simply integrated the second term using the usual properties of the delta function. So we have

$$\frac{d}{dx'}\, \delta(x - x') = -\delta(x - x')\frac{d}{dx'} \tag{3.141}$$

with the same comments as before about d/dx' acting on functions of x' on its right and the expression making sense only when acting inside an integral.

Infinite-Dimensional Generalization: Properties of the Dirac Delta Function (cont.)

The change of sign between Equations 3.138 and 3.141 makes sense because of the sign difference between x and x' in the argument of the δ function. Another way of looking at it is to recognize that, for an arbitrary function $g(x)$, it holds that

$$\frac{d}{dx}g(x-x') = \frac{d}{dy}g(y)\bigg|_{y=x-x'}, \quad \frac{d}{dx}(x-x') = \frac{d}{dy}g(y)\bigg|_{y=x-x'} \tag{3.142}$$

and also

$$\frac{d}{dx'}g(x-x') = \frac{d}{dy}g(y)\bigg|_{y=x-x'}, \quad \frac{d}{dx'}(x-x') = -\frac{d}{dy}g(y)\bigg|_{y=x-x'} \tag{3.143}$$

One can pretty easily infer that the action of an arbitrary-order derivative of the delta function is similar:

$$\frac{d^n}{dx^n}\delta(x-x') = \delta(x-x')\frac{d^n}{dx'^n} \tag{3.144}$$

Alternate representation of the delta function

You learned in Ph2/12 about Fourier transforms. The Fourier transform of a function $f(x)$ is

$$\widetilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx \, f(x) \, e^{-ikx} \tag{3.145}$$

and the inverse transform of $\widetilde{f}(k)$ is

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \, \widetilde{f}(k) \, e^{ikx} \tag{3.146}$$

(By the way, $\widetilde{f}(k)$ is example of a different representation of $|f\rangle$, $\widetilde{f}(k) = \langle k|f\rangle$, as we shall see shortly.)

Putting the two definitions together gives

$$f(x) = \int_{-\infty}^{\infty} dk \, \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx' \, f(x') \, e^{ik(x-x')} \tag{3.147}$$

$$= \int_{-\infty}^{\infty} dx' \, f(x') \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} dk \, e^{ik(x-x')} \right) \tag{3.148}$$

Comparing this expression to our defining relation for the delta function immediately implies

$$\delta(x - x') = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \, e^{ik(x-x')} \tag{3.149}$$

We will find this expression for the delta function useful later.

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

$$|x'\rangle \longleftrightarrow \delta(x - x') \tag{3.150}$$

But, is there an operator whose eigenbasis is the $\{|x\rangle\}$?

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:

$$X|x\rangle = x|x\rangle \tag{3.151}$$

where x is the position along the interval to which $|x\rangle$ corresponds. That is, we define 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. In our original finite-dimensional setup, it would have been equivalent to defining an operator *X* that just returned x_i for the ket $|i\rangle$.

Infinite-Dimensional Generalization: The *X* Operator and the {|*x*⟩} 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

$$X_{x'x} \equiv \langle x' | X | x \rangle = x \langle x' | x \rangle = x \, \delta(x' - x) \tag{3.152}$$

(We introduce the $_{x'x}$ 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

$$|g\rangle = X |f\rangle \tag{3.153}$$

Let's expand $|g\rangle$ in our $|x\rangle$ basis:

$$\langle x | g \rangle = \langle x | X | f \rangle = \int_a^b dx' \langle x | X | x' \rangle \langle x' | f \rangle = \int_a^b dx' x' \delta(x - x') f(x') = x \, f(x) \tag{3.154}$$

So, $|g\rangle = |x f\rangle$ or $g(x) = x \, f(x)$.

Lecture 8:
Infinite-Dimensional Generalization continued

Revision Date: 2007/10/18

Infinite-Dimensional Generalization: The *K* Operator and its Eigenbasis

Making the Derivative Operator Hermitian

We know that taking derivatives converts functions to other functions. Let's make this an operator *D* on our space by defining

$$D \,|f\rangle = |df/dx\rangle \tag{3.155}$$

We are implicitly making this definition in the $\{|x\rangle\}$ basis, but that is sensible since we have no other basis yet for our space! The matrix elements of *D* are easy to find. We recognize from our initial rules for the infinite extension that

$$\langle x \,|\, D \,|f\rangle = \left\langle x \,\middle|\, \frac{df}{dx} \right\rangle = \frac{df}{dx} \tag{3.156}$$

We also may write

$$\langle x \,|\, D \,|f\rangle = \int_a^b dx'\langle x\,|\,D\,|x'\rangle\langle x'\,|f\rangle = \int_a^b dx'\langle x\,|\,D\,|x'\rangle f(x') \tag{3.157}$$

Putting the two together gives

$$\int_a^b dx' \langle x \,|\, D \,|\, x' \rangle f(x') = \frac{d}{dx} f(x) \tag{3.158}$$

That is, $\langle x \,|\, D \,|\, x' \rangle$ has the same behavior as $\left(\frac{d}{dx} \delta(x - x') \right) = \delta(x - x') \frac{d}{dx'}$; *i.e.*,

$$D_{xx'} \equiv \langle x \,|\, D \,|\, x' \rangle = \left(\frac{d}{dx} \delta(x - x') \right) = \delta(x - x') \frac{d}{dx'} \tag{3.159}$$

Unfortunately, *D* is not Hermitian. The conjugate transpose of the above matrix element is obtained by exchanging $x \leftrightarrow x'$; the conjugation does nothing because everything is real. So we have

$$D^*_{x'x} = \frac{d}{dx'} \delta(x' - x) = \frac{d}{dx'} \delta(x - x') = -\frac{d}{dx} \delta(x - x') = -D_{xx'} \tag{3.160}$$

where the first step is obtained by taking the conjugate transpose, the second by using the evenness of the delta function, and the third using Equations 3.138 and 3.141 together. So we have that *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

$$K = -iD \tag{3.161}$$

(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

$$\langle g \,|\, K \,|\, f \,\rangle = \langle g \,|\, K f \,\rangle = \langle K f \,|\, g \,\rangle^* = \langle f \,|\, K^\dagger \,|\, g \,\rangle^* = \langle f \,|\, K \,|\, g \,\rangle^* \tag{3.162}$$

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:

$$\langle g\,|\,K\,|f\,\rangle = \int_a^b dx \int_a^b dx'\,\langle g\,|x\rangle\langle x\,|\,K\,|x'\rangle\langle x'\,|f\,\rangle \tag{3.163}$$

$$= \int_a^b dx\,g^*(x)\left[-i\frac{df}{dx}\right] = -i\int_a^b dx\,g^*(x)\left[\frac{df}{dx}\right]$$

$$\langle f\,|\,K\,|g\,\rangle^* = \left[\int_a^b dx \int_a^b dx'\,\langle f\,|x\rangle\langle x\,|\,K\,|x'\rangle\langle x'\,|g\,\rangle\right]^* \tag{3.164}$$

$$= \left[\int_a^b dx\,f^*(x)\left[-i\frac{dg}{dx}\right]\right]^* = i\int_a^b dx\,\left[\frac{dg^*}{dx}\right]f(x)$$

These two expressions are equal via integration by parts only if the surface term vanishes:

$$-i\,g^*(x)\,f(x)\big|_a^b \tag{3.165}$$

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.**

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 9:
Infinite-Dimensional Generalization continued

Revision Date: 2007/10/21

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. Let us denote the eigenvalues and eigenvectors as $\{k\}$ and $\{|k\rangle\}$. We of course require

$$K|k\rangle = k|k\rangle \tag{3.166}$$

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:

$$\langle x \,|\, K \,|\, k \rangle = k \, \langle x \,|\, k \rangle \tag{3.167}$$

$$\int_a^b dx\,'\langle x \,|\, K \,|\, x\,'\rangle\langle x\,'\,|\,k\rangle = k\,\psi_k(x) \tag{3.168}$$

$$-i\int_a^b dx\,'\delta(x-x\,')\frac{d}{dx\,'}\psi_k(x) = k\,\psi_k(x) \tag{3.169}$$

$$-i\,\frac{d}{dx}\psi_k(x) = k\,\psi_k(x) \tag{3.170}$$

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.

We have a simple differential equation defining $\psi_k(x)$; the solution is

$$\langle x \,|\, k \rangle = \psi_k(x) = A\,e^{ikx} \tag{3.171}$$

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. For simplicity, let's assume the interval is $[0, L]$; any other endpoints can be obtained by translation and setting $L = b - a$. The condition on the endpoints means that we may only consider solutions of the form

$$\psi_k(x) = A \cos(k x) + B \sin(k x) \tag{3.172}$$

A and B real

$$k = \frac{2m}{2} \frac{2\pi}{L} m \quad l = 1, 2, \ldots \text{ if } A \text{ and } B \text{ are both nonzero}$$

$$k = \frac{2m + 1}{2} \frac{2\pi}{L} m \quad l = 1, 2, \ldots \text{ if } A = 0 \text{ and } B \text{ is nonzero}$$

so that the function is equal at the endpoints. k is discretized, though has (countably) infinitely many allowed values. Cosines and sines are allow when an integral number of wavelengths fits in L; only sines are allows if a half-integral number of wavelengths fits in L. The natural normalization is $A = \sqrt{2/L}$ so that $\langle k | k \rangle = 1$ (you can check the integral using trigonometric identities).

▶ **Infinite interval** $(-\infty, \infty)$

Depending on whether we are consider real or complex functions, our solution is

$$\psi_k(x) = A \cos kx + B \sin kx \quad \text{with A, B, and } k \text{ any real numbers} \tag{3.173}$$

$$\psi_k(x) = \alpha \, e^{ikx} \quad \text{with } \alpha \text{ and complex and } k \text{ any real number} \tag{3.174}$$

because there is now no condition on k (note that the two solutions have the same number of coefficient degrees of freedom). The surface term is problematic in that its value is ill-defined: cos, sin, and e^{ikx} take 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

$$
\begin{aligned}
e^{i(k-k')x}\Big|_{-\infty}^{\infty} &= \left[\lim_{\beta \to 0} e^{-\beta|x|} e^{i(k-k')x} \right]\Big|_{-\infty}^{\infty} \\
&= \lim_{\beta \to 0} \left[e^{-\beta|x|} e^{i(k-k')x} \right]\Big|_{-\infty}^{\infty} \\
&= \lim_{\beta \to 0} 0 = 0
\end{aligned}
\tag{3.175}
$$

Exchanging the limit in β and the implied limit of the endpoints going to $\pm\infty$ is not a mathematically rigorous procedure, but it turns out to be correct. We will employ this somewhat dubious limiting procedure not infrequently in this course.

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, which can be obtained by setting $A = 1/\sqrt{2\pi}$:

$$\langle k\,|\,k'\rangle = \int_{-\infty}^{\infty} dx \langle k\,|x\rangle\langle x\,|k'\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\, e^{i(k'-k)x} \qquad (3.176)$$
$$= \delta(k'-k) = \delta(k-k')$$

where we have implicitly used the same type of limiting procedure as above.

Infinite-Dimensional Generalization: The K Operator and its Eigenbasis (cont.)

We note another complication arising from our transition to infinite dimensions. We proved earlier that the eigenvalues of a Hermitian operator in finite dimensions are real. However, it is certainly true that $k = k_1 + ik_2$ where k_1 and k_2 are real is a valid solution to the differential equation for ψ_k and k. Our usual proof of the realness of the eigenvalues is

$$(k_1 + ik_2)\langle k\,|k\,\rangle = k\langle k\,|k\,\rangle = \langle k\,|\,K\,|k\,\rangle = \langle k\,|\,K^\dagger|k\,\rangle \qquad (3.177)$$
$$= k^*\langle k\,|k\,\rangle = (k_1 - ik_2)\langle k\,|k\,\rangle \qquad (3.178)$$

The reason the proof does not go through is that $K = K^\dagger$ only holds when the surface term vanishes. If k has an imaginary part, that results in an exponential with real argument that blows up at either ∞ or $-\infty$. The surface term therefore does not vanish. If we restrict our space to functions that are either normalizable to unity or to a δ function – *i.e.*, functions that either behave like e^{ikx} at $\pm\infty$ or that vanish at both $\pm\infty$ – then we eliminate the complex eigenvalues that would have ruined Hermiticity. The restricted space is the physical inner product space or physical Hilbert space.

Infinite-Dimensional Generalization: The *K* Operator and its Eigenbasis (cont.)

Expansion of Vector Space Elements in the *K* Eigenbasis

We assume that our finite-dimensional proof that the eigenbasis of a Hermitian operator is an orthonormal basis for the space continues to hold, with the modification to include δ function normalization. We may thus conclude that the $\{|k\rangle\}$ eigenbasis is **complete**; *i.e.*, that

$$\sum_{j=1}^{\infty} |k_j\rangle\langle k_j| = I \quad \text{or} \quad \int_{-\infty}^{\infty} dk\, |k\rangle\langle k| = I \tag{3.179}$$

depending on whether we consider a finite or infinite interval. With the above holding, we may expand any ket in terms of the $\{|k\rangle\}$ basis. We have

$$|f\rangle = \sum_{j=1}^{\infty} |k_j\rangle\langle k_j|f\rangle \quad \text{or} \quad |f\rangle = \int_{-\infty}^{\infty} dk\, |k\rangle\langle k|f\rangle \tag{3.180}$$

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):

$$\langle k_j | f \rangle = \int_0^L dx \, \langle k_j | x \rangle \langle x | f \rangle = \sqrt{\frac{2}{L}} \int_0^L dx \, e^{-ik_j x} f(x) \tag{3.181}$$

$$\langle k | f \rangle = \int_{-\infty}^{\infty} dx \, \langle k | x \rangle \langle x | f \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx \, e^{-ikx} f(x) \tag{3.182}$$

These forms should look familiar – we see that $\langle k | f \rangle$ is just the Fourier transform of $f(x)$, which we denote by $\widetilde{f}(k)$. 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: the default basis is the $\{|x\rangle\}$ basis, and the coefficients of the expansion in this basis are $\langle x | f \rangle = f(x)$; but one can also expand $|f\rangle$ in the $\{|k\rangle\}$ basis, and the coefficients of that expansion are $\langle k | f \rangle = \widetilde{f}(k)$.

X and *K* in the *K* Eigenbasis

We have defined *X* and *K* in the *X* eigenbasis; their matrix elements are

$$\langle x \, | \, X | x' \rangle = x \, \delta(x - x') \tag{3.183}$$

$$\langle x \, | \, K \, | x' \rangle = -i \left[\frac{d}{dx} \delta(x - x') \right] = -i \, \delta(x - x') \frac{d}{dx'} \tag{3.184}$$

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

$$\langle k \, | \, K \, | k' \rangle = k \, \delta(k - k') \tag{3.185}$$

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:

$$\langle k \,|\, X \,|\, k' \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' \langle k \,|\, x \rangle \langle x \,|\, X \,|\, x' \rangle \langle x' \,|\, k' \rangle \tag{3.186}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' \, e^{-ikx} x \, \delta(x - x') e^{ik'x'} \tag{3.187}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \, x \, e^{-i(k-k')x} = i \frac{d}{dk} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} dx \, e^{-i(k-k')x} \right] \tag{3.188}$$

$$= i \frac{d}{dk} \delta(k - k') = i \, \delta(k - k') \frac{d}{dk'} \tag{3.189}$$

Hence, the action of *X* on a ket $|f\rangle$ is

$$\langle k \,|\, X \,|\, f \rangle = i \frac{d\widetilde{f}(k)}{dk} \quad \text{or, somewhat misleadingly,} \quad X|f\rangle = \left| i \frac{d\widetilde{f}(k)}{dk} \right\rangle \tag{3.190}$$

(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 $[X, K]$:

$$\langle x \,|\, X\,K \,|\, f \rangle = \int_{-\infty}^{\infty} dx' \langle x \,|\, X | x' \rangle \langle x' \,|\, K \,|\, f \rangle = \int_{-\infty}^{\infty} dx' x \, \delta(x - x') \, (-i) \, \frac{df}{dx'} \tag{3.191}$$

$$= -i \, x \, \frac{df}{dx} \tag{3.192}$$

$$\langle x \,|\, K\,X \,|\, f \rangle = \int_{-\infty}^{\infty} dx' \langle x \,|\, K | x' \rangle \langle x' \,|\, X \,|\, f \rangle = \int_{-\infty}^{\infty} dx' (-i) \, \delta(x - x') \frac{d}{dx'} x' \, f(x') \tag{3.193}$$

$$= -i \int_{-\infty}^{\infty} dx' \, \delta(x - x') \left[f(x') + x' \frac{df(x')}{dx'} \right] = -i \, f(x) - i \, x \frac{df}{dx} \tag{3.194}$$

$$\neq \langle x \,|\, X\,K \,|\, f \rangle \tag{3.195}$$

$$\implies \quad \langle x \,|\, [X, K] \,|\, f \rangle = i \, f(x) = i \, \langle x \,|\, f \rangle \quad \Longleftrightarrow \quad [X, K] = i \, I \tag{3.196}$$

Section 4
Postulates Revisited

Lecture 10:
Postulates of Quantum Mechanics Revisited

Revision Date: 2007/10/22

Summary

Recall the Postulates we briefly discussed in Section 1.2:

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.

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 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 a set of objects that 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.
- ▶ At least one orthonormal set of basis states can be defined and all states written as linear combinations of them.
- ▶ In terms of any particular basis, the state can be written as a column or row vector.
- ▶ Operators can act on the states and return new states, and an operator has a matrix representation for any particular choice of basis.
- ▶ There are Hermitian operators that have real eigenvalues and a set of eigenvectors that can be used as an orthonormal basis.
- ▶ There are unitary operators that can be used to rotate from orthonormal basis to another.

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

$$|\chi\rangle = \frac{\alpha|\psi\rangle + \beta|\phi\rangle}{\sqrt{|\alpha|^2 + |\beta|^2}} \tag{4.1}$$

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 Hilbert space of states, with X and P having the following matrix elements when using the position basis for the Hilbert space:

$$\langle x | X | x' \rangle = x \delta \left(x - x' \right) \qquad \langle x | P | x' \rangle = -i\hbar \frac{d}{dx} \delta \left(x - x' \right) \qquad (4.2)$$

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

$$\Omega(X, P) = \omega(x \to X, p \to P) \qquad (4.3)$$

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 xp, should one use XP, PX, their sum, or their difference?).

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 that we must take as a postulate for now. It can never be explicitly proven, but it can be motivated.

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.

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 $\{|\omega\rangle\}$.

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

$$P(\omega) = |P_\omega |\psi \,\rangle|^2 = \langle \psi \,|\, P_\omega \, P_\omega |\psi \,\rangle = \langle \psi \,|\, P_\omega |\psi \,\rangle \tag{4.4}$$

where P_ω is the projection operator for the ω subspace,

$$P_\omega = \sum_{\omega_i = \omega} |\omega_i \,\rangle\langle \omega_i \,| \tag{4.5}$$

and where we have written out three equivalent expressions using the fact that the projection operator P_ω is Hermitian and satisfies $P_\omega^2 = P_\omega$. In the absence of degeneracy, the above generalization simplifies to our original postulate because $P_\omega = |\omega \,\rangle\langle \omega \,|$, so

$$|\langle \omega \,|\psi \,\rangle|^2 = \langle \psi \,|\omega \,\rangle\langle \omega \,|\psi \,\rangle = \langle \psi \,|\, P_\omega |\psi \,\rangle = \langle \psi \,|\, P_\omega P_\omega |\psi \,\rangle = |P_\omega |\psi \,\rangle|^2 \tag{4.6}$$

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

$$P(\omega_i) = \frac{|\langle \omega_i | \psi \rangle|^2}{\sum_{j=1}^{n} |\langle \omega_j | \psi \rangle|^2} \tag{4.7}$$

In fact, for a properly normalized state, the denominator is trivial:

$$\sum_{j=1}^{n} |\langle \omega_j | \psi \rangle|^2 = \sum_{j=1}^{n} \langle \psi | \omega_j \rangle \langle \omega_j | \psi \rangle = \langle \psi | \psi \rangle = 1 \tag{4.8}$$

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 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

$$P(\omega)\, d\omega = \frac{|\langle \omega\,|\psi\,\rangle|^2 d\omega}{\int_{\omega_-}^{\omega_+} |\langle \omega\,|\psi\,\rangle|^2 d\omega} \tag{4.9}$$

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

$$P(\omega_1 < \omega < \omega_2) = \frac{\int_{\omega_1}^{\omega_2} d\omega\, |\langle \omega\,|\psi\,\rangle|^2}{\int_{\omega_-}^{\omega_+} d\omega\, |\langle \omega\,|\psi\,\rangle|^2} \tag{4.10}$$

These formulae can become problematic in the case of delta-function normalized states – they will always vanish because the denominator is infinite. **Relative** probabilities are still well-defined; and we shall see that, if we ask about absolute probabilities in such cases (which we will rarely do), we will be ready for probabilities that always vanish. Note that we can create states that are reasonably normalized – these simply will not be eigenstates of X or P.

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

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 $\widetilde{\psi}(k) = \langle k | \psi \rangle$ 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 $\widetilde{\psi}(k)$ are in general very different functions – so, the argument of the function, which is normally a dummy variable, means something. To be clear, we could use a labelling like $\psi_\omega(\omega) = \langle \omega | \psi \rangle$ so that $\psi_x(x) = \langle x | \psi \rangle$ and $\psi_k(k) = \langle k | \psi \rangle$.

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 $\{|i\rangle\}$ that have eigenvalues $\{\omega_i\}$ and $\{\lambda_i\}$. If $|\psi\rangle$ is an eigenstate $|i\rangle$, then measurements of Ω and Λ will yield the definite values ω_i and λ_i. If $|\psi\rangle$ is not an eigenstate, then the measurement outcomes will be correlated: if Ω yields ω_i, then Λ yields λ_i because the projection operator $P_{\omega=\omega_i}$ is the same as the projection operator $P_{\lambda=\lambda_i i}$. The relative probabilities $P(\omega_i)$ and $P(\lambda_i)$ 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.

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

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,

$$\langle \Omega \rangle = \sum_i P(\omega_i)\, \omega_i \quad \text{or} \quad \langle \Omega \rangle = \int_{\omega_-}^{\omega_+} d\omega\, P(\omega)\, \omega \qquad (4.11)$$

We can write this explicitly in terms of Ω and the state $|\psi\rangle$:

$$\langle \Omega \rangle = \sum_i |\langle \omega_i |\psi\rangle|^2 \omega_i = \sum_i \langle \psi |\omega_i\rangle\langle\omega_i |\psi\rangle \omega_i = \sum_i \langle \psi |\Omega|\omega_i\rangle\langle\omega_i |\psi\rangle \qquad (4.12)$$

$$= \langle \psi |\Omega|\psi\rangle$$

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

$$\langle (\Delta\Omega)^2 \rangle = \sum_i P(\omega_i) \left(\omega_i - \langle\Omega\rangle\right)^2 \quad \text{or} \quad \langle (\Delta\Omega)^2 \rangle = \int_{\omega_-}^{\omega_+} d\omega \, P(\omega) \left(\omega - \langle\Omega\rangle\right)^2 \quad (4.13)$$

Let's pursue this in the discretized case. First, the above expression can be simplified:

$$\sum_i P(\omega_i) \left(\omega_i - \langle\Omega\rangle\right)^2 = \sum_i P(\omega_i)\omega_i^2 - 2\langle\Omega\rangle \sum_i P(\omega_i)\omega_i + \langle\Omega\rangle^2 \sum_i P(\omega_i) \quad (4.14)$$

$$= \left[\sum_i P(\omega_i) \, \omega_i^2 \right] - \langle\Omega\rangle^2$$

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.

Let's write out the first term in terms of Ω and $|\psi\rangle$.

$$\sum_i P(\omega_i)\,\omega_i^2 = \sum_i \langle\psi\,|\omega_i\,\rangle\langle\omega_i\,|\psi\,\rangle\,\omega_i^2 = \sum_i \langle\psi\,|\Omega|\omega_i\,\rangle\langle\omega_i\,|\Omega|\psi\rangle = \langle\psi\,|\Omega^2|\psi\rangle \quad (4.15)$$

where we again used completeness. So we have

$$\langle(\Delta\Omega)^2\rangle = \langle\psi\,|\Omega^2|\psi\rangle - \langle\Omega\rangle^2 = \langle\psi\,|\left[\Omega^2 - \langle\Omega\rangle^2\right]|\psi\rangle = \langle\psi\,|\,[\Omega - \langle\Omega\rangle]^2\,|\psi\rangle \quad (4.16)$$

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.

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

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle \qquad (4.17)$$

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

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, but we will delay them until after you have encountered Hamiltonian mechanics in Ph106a.

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:

▶ **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

$$i\,\hbar\,\frac{dU(t)}{dt}\,|\psi(0)\rangle = H\,U(t)\,|\psi(0)\rangle \tag{4.18}$$

$|\psi(0)\rangle$ is constant, so this is really just a differential equation for $U(t)$:

$$i\,\hbar\,\frac{dU(t)}{dt} = H\,U(t) \tag{4.19}$$

Our discussion of calculus with operators (Section 3.8) tells us the solution is

$$U(t) = e^{-\frac{i}{\hbar}\,H\,t} \tag{4.20}$$

Thus, our full solution for the time evolution of the state is

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle = e^{-\frac{i}{\hbar}\,H\,t}|\psi(0)\rangle \tag{4.21}$$

Since we assume H is Hermitian, we are guaranteed the propagator $U(t)$ is
unitary and thus preserves the norm $|\psi|^2$.

▶ **Eigenbasis form**
   H is a Hermitian operator, so it has eigenvalues $\{E_i\}$ (discrete or continuous). If we consider an eigenstate $|E_i\rangle$, then we have

$$H|E_i\rangle = E_i|E_i\rangle \tag{4.22}$$

for all time (assuming time-independent H). Thus, for this state, our differential equation becomes

$$i\hbar \frac{d}{dt}|E_i(t)\rangle = H|E_i(t)\rangle = E_i|E_i(t)\rangle \tag{4.23}$$

Since $|E_i(t)\rangle$ has to remain an eigenvector of H, only its coefficient may change. The differential equation suggests the solution $|E_i(t)\rangle = e^{-\frac{i}{\hbar}E_i t}|E_i(0)\rangle$; one can check trivially that it satisfies the equation.

With this in hand, we can calculate the time evolution of any state:

$$|\psi(t)\rangle = \sum_i |E_i(t)\rangle\langle E_i(t)|\psi(t)\rangle \tag{4.24}$$

This does not get us very far yet, but let's note something useful:

$$\frac{d}{dt}\langle E_i(t)|\psi(t)\rangle = \left[\frac{d}{dt}\langle E_i(t)|\right]|\psi(t)\rangle + \langle E_i(t)|\left[\frac{d}{dt}|\psi(t)\rangle\right] \tag{4.25}$$

$$= \left[\frac{d}{dt}|E_i(t)\rangle\right]^\dagger |\psi(t)\rangle + \langle E_i(t)|\left[-\frac{i}{\hbar}H|\psi(t)\rangle\right] \tag{4.26}$$

$$= \left[-\frac{i}{\hbar}H|E_i(t)\rangle\right]^\dagger |\psi(t)\rangle - \frac{i}{\hbar}\langle E_i(t)|H|\psi(t)\rangle \tag{4.27}$$

$$= \left[-\frac{i}{\hbar}E_i|E_i(t)\rangle\right]^\dagger |\psi(t)\rangle - \frac{i}{\hbar}[H|E_i(t)\rangle]^\dagger |\psi(t)\rangle \tag{4.28}$$

$$= \frac{i}{\hbar}E_i\langle E_i(t)|\psi(t)\rangle - \frac{i}{\hbar}[E_i|E_i(t)\rangle]^\dagger |\psi(t)\rangle \tag{4.29}$$

$$= \frac{i}{\hbar}E_i\langle E_i(t)|\psi(t)\rangle - \frac{i}{\hbar}E_i\langle E_i(t)|\psi(t)\rangle = 0 \tag{4.30}$$

We see that the projection of $|\psi(t)\rangle$ onto $|E_i(t)\rangle$ 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 $\langle E_i(t)|\psi(t)\rangle = \langle E_i(0)|\psi(0)\rangle$ and thus

$$|\psi(t)\rangle = \sum_i |E_i(t)\rangle\langle E_i(0)|\psi(0)\rangle = \sum_i e^{-\frac{i}{\hbar}E_i t}|E_i(0)\rangle\langle E_i(0)|\psi(0)\rangle \quad (4.31)$$

Combining the two forms, we have an explicit form for the propagator $U(t)$:

$$U(t) = e^{-\frac{i}{\hbar}Ht} = \sum_i e^{-\frac{i}{\hbar}E_i t}|E_i(0)\rangle\langle E_i(0)| = \sum_i e^{-\frac{i}{\hbar}E_i t}P_{E_i(0)} \quad (4.32)$$

where $P_{E_i(0)}$ is the projection operator onto the eigenvector $|E_i(0)\rangle$. The similarity to the unitary propagator we found for Example 4, Equation 3.88, should be clear.

We will in general drop the (0) in $|E_i(0)\rangle$ and $\langle E_i(0)|$ and assume that, when we indicate the eigenvectors of the Hamiltonian $|E_i\rangle$, we mean constant vectors that do not include the unitary time evolution.

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 4.19, now becomes

$$i\hbar \frac{dU(t)}{dt} = H(t)\, U(t) \tag{4.33}$$

The solution is no longer $U(t) = e^{-\frac{i}{\hbar} H t}$: if one takes the time derivative, one gets

$$i\hbar \frac{dU(t)}{dt} = \left[H(t) + t\, \frac{dH(t)}{dt} \right] U(t) \tag{4.34}$$

Rather, as Shankar shows, the solution becomes

$$U(t) = \mathcal{T}\left[\exp\left(-\frac{i}{\hbar} \int_0^t dt'\, H(t') \right) \right] = \lim_{N \to \infty} \prod_{j=0}^{N-1} \exp\left[-\frac{i}{\hbar} \left(\frac{t}{N} \right) H\left(\frac{j}{N}\, t \right) \right] \tag{4.35}$$

where $\mathcal{T}[\ ]$ denotes the time-ordered integral. That is, one uses the standard solution for infinitesimally small time-evolutions and then takes the product of them all. Since we will not encounter such time-dependent Hamiltonians regularly, we neglect derivation or discussion of this result.

Section 5
Simple One-Dimensional Problems

Lecture 11:
The One-Dimensional Free Particle

Revision Date: 2007/11/02

The Free Particle

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 Equations 4.31 and 4.32, reproduced here:

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle \qquad U(t) = e^{-\frac{i}{\hbar}Ht} = \sum_i e^{-\frac{i}{\hbar}E_i t}|E_i(0)\rangle\langle E_i(0)|$$

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.) The eigenvalues of H are given by acting on these eigenstates with H:

$$H|p\rangle = \frac{P^2}{2m}|p\rangle = \frac{p^2}{2m}|p\rangle \equiv E|p\rangle \tag{5.2}$$

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:

$$U(t) = \int_{-\infty}^{\infty} dp \, e^{-\frac{i}{\hbar} \frac{p^2}{2m} t} |p\rangle\langle p| \tag{5.4}$$

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 4.32.

We note as an aside that the above is **not** the same as

$$U(t) = \int_{0}^{\infty} dE \, e^{-\frac{i}{\hbar} E t} \left(|E, +\rangle\langle E, +| + |E, -\rangle\langle E, -| \right) \tag{5.5}$$

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 $E = p^2/2m$, 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:

$$[U(t)]_{xx'} = \langle x | U(t) | x' \rangle = \int_{-\infty}^{\infty} dp \, e^{-\frac{i}{\hbar} \frac{p^2}{2m} t} \langle x | p \rangle \langle p | x' \rangle \tag{5.6}$$

$$= \int_{-\infty}^{\infty} dp \, e^{-\frac{i}{\hbar} \frac{p^2}{2m} t} e^{\frac{i}{\hbar} p(x-x')} \tag{5.7}$$

$$= \sqrt{\frac{m}{2 \pi \hbar i t}} \, e^{\frac{i}{\hbar} \frac{m(x-x')^2}{2t}} \tag{5.8}$$

Proving this 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

$$\left[U(t'-t)\right]_{xx'} = \sqrt{\frac{m}{2\pi\hbar i(t'-t)}}\, e^{\frac{i}{\hbar}\frac{(x-x')^2}{2(t'-t)}} \tag{5.9}$$

The Free Particle (cont.)

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 *wavefunction*) one gets at time t if one's initial state is $|x'\rangle$. That state has $\{|x\rangle\}$-basis representation $\langle x|x'\rangle = \delta(x - x')$. That is: let $|\psi(0)\rangle = |x'\rangle$. Then

$$\langle x|\psi(t)\rangle = \langle x|U(t)|\psi(0)\rangle = \int_{-\infty}^{\infty} dx'' \langle x|U(t)|x''\rangle \langle x''|\psi(0)\rangle \tag{5.10}$$

$$= \int_{-\infty}^{\infty} dx'' \sqrt{\frac{m}{2\pi\hbar i t}} \, e^{\frac{i}{\hbar}\frac{m(x-x'')^2}{2t}} \delta(x'' - x') \tag{5.11}$$

$$= \sqrt{\frac{m}{2\pi\hbar i t}} \, e^{\frac{i}{\hbar}\frac{m(x-x')^2}{2t}} = [U(t)]_{xx'} \tag{5.12}$$

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 $|[U(t)]_{xx'}|^2$ is the probability that a particle that is perfectly localized to x' at $t = 0$ will be detected at x at time t. More generally, U tells us how the probability of finding a particle at position x at time t is determined by the initial wavefunction $\langle x|\psi(0)\rangle$. 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).

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

$$\langle x|\psi(0)\rangle = \left(\frac{1}{2\pi\sigma_x^2}\right)^{1/4} e^{\frac{i}{\hbar}p_0 x} e^{-\frac{x^2}{4\sigma_x^2}} \tag{5.13}$$

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 $\langle(\Delta X)^2\rangle = \sigma_x^2$.

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:

$$\langle X\rangle_{t=0} = \langle\psi(0)\,|X\,|\psi(0)\,\rangle = \int_{-\infty}^{\infty} dx\,\langle\psi(0)\,|X\,|x\rangle\langle x\,|\psi(0)\,\rangle \tag{5.14}$$

$$= \int_{-\infty}^{\infty} dx\,\langle\psi(0)\,|x\rangle\,x\,\langle x\,|\psi(0)\,\rangle \tag{5.15}$$

$$= \int_{-\infty}^{\infty} dx\,x\left(\frac{1}{2\,\pi\,\sigma_x^2}\right)^{1/2} e^{-\frac{x^2}{2\,\sigma_x^2}} = 0 \tag{5.16}$$

Second, we can calculate the initial variance $\langle(\Delta X)^2\rangle_{t=0}$:

$$\langle(\Delta X)^2\rangle_{t=0} = \int_{-\infty}^{\infty} dx\left(x^2 - \langle X\rangle_{t=0}^2\right)\left(\frac{1}{2\,\pi\,\sigma_x^2}\right)^{1/2} e^{-\frac{x^2}{2\,\sigma_x^2}} = \sigma_x^2 \tag{5.17}$$

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):

$$\langle p\,|\psi(0)\,\rangle = \int_{-\infty}^{\infty} dx\,\langle p\,|x\rangle\langle x\,|\psi(0)\,\rangle \tag{5.18}$$

$$= \frac{1}{\sqrt{2\,\pi\,\hbar}}\int_{-\infty}^{\infty} dx\,e^{-\frac{i}{\hbar}\,p\,x}\left(\frac{1}{2\,\pi\,\sigma_x^2}\right)^{1/4}e^{\frac{i}{\hbar}\,p_0\,x}e^{-\frac{x^2}{4\,\sigma_x^2}} \tag{5.19}$$

$$= \left(\frac{1}{2\,\pi\,\sigma_p^2}\right)^{1/4}e^{-\frac{(p-p_0)^2}{4\,\sigma_p^2}} \tag{5.20}$$

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

$$\langle P\rangle_{t=0} = p_0 \qquad \langle(\Delta P)^2\rangle_{t=0} = \sigma_p^2 \tag{5.21}$$

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

$$\langle p|\psi(t)\rangle = \left(\frac{1}{2\pi\sigma_p^2}\right)^{1/4} e^{-\frac{(p-p_0)^2}{4\sigma_p^2}} e^{-\frac{i}{\hbar}\frac{p^2}{2m}t} \tag{5.22}$$

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

$$\langle P\rangle_t = \langle P\rangle_{t=0} = p_0 \qquad \langle(\Delta P)^2\rangle_t = \langle(\Delta P)^2\rangle_{t=0} = \sigma_p^2 \tag{5.23}$$

The Free Particle (cont.)

Let's also calculate the $\{|x\rangle\}$ representation of $|\psi(t)\rangle$. Here, we can just use our propagator formula, Equation 5.8, which tells us

$$\langle x | \psi(t) \rangle = \int_{-\infty}^{\infty} dx' \, [U(t)]_{xx'} \, \langle x' | \psi(0) \rangle \tag{5.24}$$

$$= \int_{-\infty}^{\infty} dx' \sqrt{\frac{m}{2\pi\hbar i t}} \, e^{\frac{i}{\hbar} \frac{m(x-x')^2}{2t}} \left(\frac{1}{2\pi\sigma_x^2} \right)^{1/4} e^{\frac{i}{\hbar} p_0 x'} e^{-\frac{(x')^2}{4\sigma_x^2}} \tag{5.25}$$

$$= \left[\frac{1}{2\pi\sigma_x^2 \left(1 + \frac{i\hbar t}{2m\sigma_x^2} \right)} \right]^{-1/4} \exp\left[-\frac{\left(x - \frac{p_0}{m} t \right)^2}{4\sigma_x^2 \left(1 + \frac{i\hbar t}{2m\sigma_x^2} \right)} \right] \tag{5.26}$$

$$\exp\left(\frac{i}{\hbar} p_0 x \right) \exp\left(-\frac{i}{\hbar} \frac{p_0^2}{2m} t \right)$$

where we do the integral in the usual fashion, by completing the square and using the Gaussian definite integral.

The probability density in the $\{|x\rangle\}$ basis is

$$|\langle x | \psi(t) \rangle|^2 = \left[\frac{1}{2\pi\sigma_x^2 \left(1 + \left(\frac{\hbar t}{2 m \sigma_x^2} \right)^2 \right)} \right]^{-1/2} \exp\left[-\frac{\left(x - \frac{p_0}{m} t \right)^2}{2\sigma_x^2 \left(1 + \left(\frac{\hbar t}{2 m \sigma_x^2} \right)^2 \right)} \right]$$

(5.27)

Because the probability density is symmetric about $x = \frac{p_0}{m} t$, it is easy to see that

$$\langle X \rangle_t = \frac{p_0}{m} t = \langle X \rangle_{t=0} + \frac{p_0}{m} t$$

(5.28)

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),

$$\langle (\Delta X)^2 \rangle_t = \sigma_x^2 \left[1 + \left(\frac{\hbar t}{2 m \sigma_x^2} \right)^2 \right] = \langle (\Delta X)^2 \rangle_{t=0} \left[1 + \left(\frac{\hbar t}{2 m \sigma_x^2} \right)^2 \right] \qquad (5.29)$$

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:

$$t \gg t_0 = 2 m \frac{\sigma_x^2}{\hbar} = m \frac{\sigma_x}{\sigma_p} = \frac{\sigma_x}{\sigma_v} \qquad (5.30)$$

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 such a use in the particle in a box problem.

The Free Particle (cont.)

Position-Momentum Uncertainty Relation

Before leaving the free particle, we note and 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 3.149, 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

$$\langle (\Delta X)^2 \rangle_{t=0} = \sigma_x^2 \qquad \langle (\Delta P)^2 \rangle_{t=0} = \sigma_p^2 = \frac{\hbar^2}{4\,\sigma_x^2} \qquad (5.31)$$

Hence, at $t = 0$, we have the uncertainty relation

$$\sqrt{\langle (\Delta X)^2 \rangle_{t=0}} \sqrt{\langle (\Delta P)^2 \rangle_{t=0}} = \frac{\hbar}{2} \qquad (5.32)$$

We saw that, for $t > 0$, the position uncertainty grows while the momentum uncertainty is unchanged, so in general we have

$$\sqrt{\langle (\Delta X)^2 \rangle} \sqrt{\langle (\Delta P)^2 \rangle} \geq \frac{\hbar}{2} \qquad (5.33)$$

We will later make a general proof of this uncertainty relationship between noncommuting observables.

Lecture 12:
The One-Dimensional Particle in a Box

Revision Date: 2007/10/27

The Particle in a Box

The Hamiltonian

A "box" consists of a region of vanishing potential energy surrounded by a region of infinite potential energy:

$$V(x) = \lim_{V_0 \to \infty} \left\{ \begin{array}{ll} 0 & |x| \leq \frac{L}{2} \\ V_0 & |x| > \frac{L}{2} \end{array} \right. \tag{5.34}$$

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

$$\mathcal{H}(x, p) = \frac{p^2}{2\,m} + V(x) \tag{5.35}$$

The Particle in a Box (cont.)

Postulate 2 tells us that the quantum Hamiltonian operator is

$$H(X, P) = \frac{P^2}{2m} + V(X) \tag{5.36}$$

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

$$\langle x | H(X, P) | x' \rangle = \langle x | \frac{P^2}{2m} | x' \rangle + \langle x | V(X) | x' \rangle \tag{5.37}$$

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:

$$\langle x | \frac{P^2}{2m} | x' \rangle = \frac{1}{2m} \int_{-\infty}^{\infty} dx'' \langle x | P | x'' \rangle \langle x'' | P | x' \rangle \tag{5.38}$$

$$= -\frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx'' \left[\frac{d}{dx} \delta(x - x'') \right] \left[\frac{d}{dx''} \delta(x'' - x') \right] \tag{5.39}$$

$$= -\frac{\hbar^2}{2m} \frac{d}{dx} \int_{-\infty}^{\infty} dx'' \delta(x - x'') \left[\frac{d}{dx''} \delta(x'' - x') \right] \tag{5.40}$$

$$= -\frac{\hbar^2}{2m} \frac{d}{dx} \left[\frac{d}{dx} \delta(x - x') \right] \tag{5.41}$$

$$= \frac{\hbar^2}{2m} \frac{d}{dx} \left[\delta(x - x') \frac{d}{dx'} \right] \tag{5.42}$$

$$= -\frac{\hbar^2}{2m} \delta(x - x') \frac{d^2}{d(x')^2} \tag{5.43}$$

where in next-to-last and last steps we used $\left[\frac{d}{dx} \delta(x - x') \right] = \delta(x - x') \frac{d}{dx'}$ twice.

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

$$\langle x | V(X) | x' \rangle = \sum_{k=0}^{\infty} V_k \langle x | X^k | x' \rangle = \sum_{k=0}^{\infty} V_k \left(x' \right)^k \langle x | x' \rangle \qquad (5.44)$$

$$= \sum_{k=0}^{\infty} V_k \left(x' \right)^k \delta(x - x') = \delta(x - x') \, V(x') \qquad (5.45)$$

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

$$\langle x | X^2 | x' \rangle = \int_{-\infty}^{\infty} dx'' \langle x | X | x'' \rangle \langle x'' | X | x' \rangle = \int_{-\infty}^{\infty} dx'' \, x \, \delta(x - x'') \, x'' \, \delta(x'' - x') \qquad (5.46)$$

$$= x^2 \delta(x - x') = (x')^2 \delta(x - x') \qquad (5.47)$$

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 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

$$\langle x \,|\, V(X) \,|\, x' \rangle = \delta(x - x') \, V(x') \tag{5.48}$$

With the above results, we have that the matrix elements of H are given by:

$$\langle x \,|\, H \,|\, x' \rangle = \delta(x - x') \left[-\frac{\hbar^2}{2\,m} \frac{d^2}{d(x')^2} + V(x') \right] \tag{5.49}$$

Thus, for an arbitrary state $|f\rangle$, we have that (using completeness as usual)

$$\langle x\,|H|f\rangle = \int_{-\infty}^{\infty} dx'\langle x\,|H|x'\rangle\langle x'\,|f\rangle \tag{5.50}$$

$$= \int_{-\infty}^{\infty} dx'\,\delta(x-x')\left[-\frac{\hbar^2}{2\,m}\frac{d^2}{d(x')^2} + V(x')\right]f(x') \tag{5.51}$$

$$= \int_{-\infty}^{\infty} dx'\,\delta(x-x')\left[-\frac{\hbar^2}{2\,m}\frac{d^2f(x')}{d(x')^2} + V(x')\,f(x')\right] \tag{5.52}$$

$$= -\frac{\hbar^2}{2\,m}\frac{d^2f(x)}{dx^2} + V(x)\,f(x) \tag{5.53}$$

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 $\langle x\,|$:

$$\langle x\,|E\,|\psi_E\rangle = \langle x\,|H\,|\psi_E\rangle = -\frac{\hbar^2}{2\,m}\,\frac{d^2}{dx^2}\,\psi_{E,x}(x) + V(x)\,\psi_{E,x}(x) \qquad (5.54)$$

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 $|\psi_E\rangle$. For the purposes of finding a mathematical solution, we rewrite:

$$\frac{d^2}{dx^2}\psi_{E,x}(x) + \frac{2\,m}{\hbar^2}\,[E - V(x)]\,\psi_{E,x}(x) = 0 \qquad (5.55)$$

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.

The Particle in a Box (cont.)

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

$$\frac{d^2}{dx^2}\psi_{E,x}(x) + \frac{2\,m}{\hbar^2}\,\alpha\,\psi_{E,x}(x) = 0 \tag{5.56}$$

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,

$$\psi_{E,x}(x) = A\,e^{-\kappa\,x} + B\,e^{\kappa\,x} \qquad \kappa = \sqrt{-\frac{2\,m}{\hbar^2}\,\alpha} \tag{5.57}$$

Let us consider states with $E < V_0$. (We will consider $E \geq V_0$ below.) For region I, the A term blows up and mst be discarded to keep the state normalizable; for region III, the B term similarly must be discarded. We therefore have

$$\psi_{E,x,I}(x) = B_I \, e^{\kappa \, x} \qquad \psi_{E,x,III}(x) = A_{III} \, e^{-\kappa \, x} \tag{5.58}$$

Next, let's consider our solution in region II. We rewrite

$$\psi_{E,x,II}(x) = A \, e^{i \, k \, x} + B \, e^{-i \, k \, x} \qquad k = \sqrt{\frac{2 \, m}{\hbar^2} \, E} \tag{5.59}$$

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:

$$\psi_{E,x}(x)\Big|_{\frac{L}{2}-\epsilon}^{\frac{L}{2}+\epsilon} = \int_{\frac{L}{2}-\epsilon}^{\frac{L}{2}+\epsilon} dx \, \frac{d}{dx} \, \psi_{E,x}(x) \tag{5.60}$$

$$= \int_{\frac{L}{2}-\epsilon}^{\frac{L}{2}} dx \, \frac{d}{dx} \, \psi_{E,x,II}(x) + \int_{\frac{L}{2}}^{\frac{L}{2}+\epsilon} dx \, \frac{d}{dx} \, \psi_{E,x,III}(x) \tag{5.61}$$

$$= \epsilon \, \frac{d}{dx} \, \psi_{E,x,II} \left(\frac{L}{2} - \epsilon \right) + \epsilon \, \frac{d}{dx} \, \psi_{E,x,III} \left(\frac{L}{2} + \epsilon \right) \tag{5.62}$$

$$= i \, k \, \epsilon \left(A_{II} \, e^{i \, k \, \frac{L}{2}} - B_{II} \, e^{-i \, k \, \frac{L}{2}} \right) - \kappa \, \epsilon \, A_{III} \, e^{-\kappa \, \frac{L}{2}} \tag{5.63}$$

Now we take the limit $V_0 \to \infty$. This has no effect on the first term. We know $e^{-\kappa \, \frac{L}{2}} \to 0$ faster than $\kappa \to \infty$ because exponentials decay faster than any polynomial. So the second term vanishes as $V_0 \to \infty$. Then 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, meaning the $\psi_{E,x} \left(\frac{L}{2} \right) = 0$ because $\psi_{E,x,III}(x) = 0$ once $\kappa \to \infty$. The same condition holds at $x = -\frac{L}{2}$ by the same argument.

Lecture 13:
The One-Dimensional Particle in a Box continued
General Considerations on Bound States
Continuity Equation for Probability

Revision Date: 2007/10/30

The Particle in a Box

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.

So, we have the following pair of equations that must be satisfied:

$$A\,e^{-i\,k\,\frac{L}{2}} + B\,e^{i\,k\,\frac{L}{2}} = 0 \tag{5.64}$$

$$A\,e^{i\,k\,\frac{L}{2}} + B\,e^{-i\,k\,\frac{L}{2}} = 0 \tag{5.65}$$

where we have dropped the $_{II}$ 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

$$\left[\begin{array}{cc} e^{-i\,k\,\frac{L}{2}} & e^{i\,k\,\frac{L}{2}} \\ e^{i\,k\,\frac{L}{2}} & e^{-i\,k\,\frac{L}{2}} \end{array}\right] \left[\begin{array}{c} A \\ B \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \end{array}\right] \tag{5.66}$$

We require the determinant vanish in order for there to be a nontrivial pair of coefficients A, B:

$$e^{-ikL} - e^{ikL} = 0 \qquad (5.67)$$

$$2i \sin(kL) = 0 \qquad (5.68)$$

which implies

$$k = \frac{n\pi}{L} \qquad n = 0, \pm1, \pm2, \ldots \qquad (5.69)$$

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 constrain on *k*:

$$A\,e^{-i\,n\,\frac{\pi}{2}} + B\,e^{i\,n\,\frac{\pi}{2}} = 0 \Longrightarrow A = -e^{i\,n\,\pi}\,B = (-1)^{n+1}B \tag{5.70}$$

The two complex exponentials are added with a $+$ or $-$ sign depending on whether *n* is odd or even, so we get a series of sines and cosines as the solutions:

$$\psi_{E,x}(x) = \sqrt{\frac{2}{L}}\,\sin\left(\frac{n\,\pi\,x}{L}\right) \qquad n > 0,\,\text{even} \tag{5.71}$$

$$\psi_{E,x}(x) = \sqrt{\frac{2}{L}}\,\cos\left(\frac{n\,\pi\,x}{L}\right) \qquad n > 0,\,\text{odd} \tag{5.72}$$

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.)

Since the interval over which the wavefunction is nonzero will always be finite, we have been able to normalization the states to unity rather than to a delta function; the norm of a state is explicitly given by

$$|\psi_E|^2 = \langle \psi_E | \psi_E \rangle = \frac{1}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} dx \, \langle \psi_E | x \rangle \langle x | \psi_E \rangle = \frac{1}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} dx \, |\psi_{E,x}(x)|^2 \qquad (5.73)$$

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 eigenvalues are now found trivially by evaluating our differential equation, Equation 5.55, in region *II* (where $V(x)$ is finite and in fact vanishes):

$$E_n \psi_{E,x}(x) = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi_{E,x}(x) = \frac{\hbar^2 k_n^2}{2m} \psi_{E,x}(x) \qquad (5.74)$$

$$\implies E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{\hbar^2 \pi^2 n^2}{2m L^2} \qquad (5.75)$$

Relation to $\{|x\rangle\}$ and $\{|p\rangle\}$ Basis States

We can 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

$$\langle x | p \rangle = \frac{1}{\sqrt{2\,\pi\,\hbar}}\, e^{\frac{i}{\hbar}\, p\, x} \tag{5.76}$$

The position representation $\psi_{E_n, x}(x) = \langle x | \psi_{E_n} \rangle$ of our eigenstate $|\psi_{E_n}\rangle$ has two differences from the above functions: : first, it is a linear combination of states $|p\rangle$ and $|-p\rangle$ with $p = \sqrt{2\,m\,E_n}$; and second, the eigenstates are only equal to the above forms for $\langle x | p \rangle$ over the interval $\left[-\frac{L}{2}, \frac{L}{2}\right]$. Hence, they simply are not momentum eigenstates.

We also make the obvious point that our energy eigenstates are not position eigenstates either: again, 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.

The Particle in a Box (cont.)

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. This consisted of four conditions: that the state vanishes for $|x| > L/2$ (two conditions) and that the wavefunction is continuous at the two boundaries. The first two conditions do not result in any kind of quantization because they could be satisfied by simply setting coefficients appropriately. But, once those two coefficients are set, there is only one coefficient degree of freedom left – the normalization of the wavefunction must always be left free by the differential equation because it is linear. So, one must use up some of the freedom in k to satisfy the fourth condition.

In the free-particle case, we have all the trappings of quantum mechanics without the quantization because of the lack of such boundary conditions.

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, but, 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,

$$\sqrt{\langle(\Delta X)^2\rangle}\sqrt{\langle(\Delta P)^2\rangle} \geq \frac{\hbar}{2} \tag{5.77}$$

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

$$\sqrt{\langle(\Delta P)^2\rangle} \geq \frac{\hbar}{2\sqrt{\langle(\Delta X)^2\rangle}} \geq \frac{\hbar}{L} \tag{5.78}$$

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 $\langle H \rangle$:

$$\langle H \rangle = \frac{\langle P^2 \rangle}{2\,m} \geq \frac{1}{2\,m} \frac{\hbar^2}{2\,\langle (\Delta X)^2 \rangle} \geq \frac{\hbar^2}{2\,m\,L^2} \tag{5.79}$$

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 4.32 to write the propagator that evolves forward any initial state:

$$U(t) = \sum_n e^{-\frac{i}{\hbar} E_n t} |\psi_{E_n}\rangle\langle\psi_{E_n}| = \sum_n \exp\left[-\frac{i}{\hbar}\frac{\hbar^2 \pi^2 n^2}{2 m L^2} t\right] |\psi_n\rangle\langle\psi_n| \tag{5.80}$$

where we use the shorthand $|\psi_n\rangle = |\psi_{E_n}\rangle$. In the $\{|x\rangle\}$-basis representation, we have

$$[U(t)]_{xx'} = \langle x|U(t)|x'\rangle = \sum_n \exp\left[-\frac{i}{\hbar}\frac{\hbar^2 \pi^2 n^2}{2 m L^2} t\right] \langle x|\psi_n\rangle\langle\psi_n|x'\rangle \tag{5.81}$$

$$= \sum_n \exp\left[-\frac{i}{\hbar}\frac{\hbar^2 \pi^2 n^2}{2 m L^2} t\right] \psi_{n,x}(x)\,\psi_{n,x}^*(x') \tag{5.82}$$

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.

General Considerations on Bound States and Quantization

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 $\epsilon \to 0$, 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).

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 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 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.

For a free state, we are allowed to keep all the coefficients (no worries about solutions blowing up at $\pm\infty$, so we really have four degrees of freedom). For free states, we expect there to be two independent states with the same energy – a right-going state and a left-going state. Thus, our four degrees of freedom break down into two degrees of freedom each for two independent solutions. These two degrees of freedom for each solution can be rewritten as phase and normalization freedom – we arbitrarily pick a phase for the solution, and we force the normalization to be something reasonable. 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.

For a bound state, the wavefunction must become real exponential at some x for a given value of E because E is less than V somewhere. This is not a condition, it is just a matter of whether we have the complex exponential or real exponential solution in that region. Once that is set, though, we acquire an additional condition we did not have in the free-particle case, that the wavefunction may not blow up at $\pm\infty$ in order for it to be normalizable. This eliminates two of 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. We only had four coefficient degrees of freedom left from the $-\infty$ interval. Two of these are always used up by phase freedom and normalization, even when there is no energy degeneracy. That leaves two coefficient degrees of freedom to satisfy four added conditions. Thus, one is forced to use up the E degree of freedom to satisfy these remaining conditions. Hence, we obtain energy quantization.

The Continuity Equation for Probability

Analogy to Electromagnetism

Postulate 3 of QM tells us to interpret $|\langle x\,|\psi(t)\,\rangle|$ as the probability $P(x)$ that the position of the particle is in the interval x to $x + dx$; $P(x)$ 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

$$Q_{dx}(x, t) = \rho(x, t)\, dx \quad \text{or} \quad Q_{d^3x}(\vec{x}, t) = \rho(\vec{x}, t)\, d^3x \tag{5.83}$$

▶ The total charge in an interval $[a, b]$ or volume V is

$$Q_{[a,b]}(t) = \int_a^b dx\, \rho(x, t) \quad \text{or} \quad Q_V(t) = \int_V d^3x\, \rho(\vec{x}, t) \tag{5.84}$$

▶ The total charge over all space is conserved.

▶ The electric current density is defined as

$$j(x, t) = \rho(x, t)\, v(x, t) \quad \text{or} \quad \vec{j}(\vec{x}, t) = \rho(\vec{x}, t)\, \vec{v}(\vec{x}, t) \tag{5.85}$$

where v or $\vec{v}$ is the velocity of the charges currently at position x or $\vec{x}$ at time t.

▶ The charge density satisfies the continuity equation

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial j(x,t)}{\partial x} = 0 \quad \text{or} \quad \frac{\partial \rho(\vec{x},t)}{\partial t} + \vec{\nabla} \cdot \vec{j}(x,t) = 0 \qquad (5.86)$$

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 charges 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 integral version of the continuity equation is

$$\frac{\partial}{\partial t} \int_a^b dx \rho(x,t) + [j(b,t) - j(a,t)] = 0 \tag{5.87}$$

$$\text{or} \quad \frac{\partial}{\partial t} \int_V d^3x \rho(\vec{x},t) + \int_{S_V} d^2x \, \hat{n}(\vec{x}) \cdot \vec{j}(\vec{x},t) = 0 \tag{5.88}$$

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.

We can prove total conservation and the continuity equation for quantum-mechanical probability density, which is

$$P(x, t) = |\langle x | \psi(t) \rangle|^2 = |\psi_x(x)|^2 \quad \text{or} \quad P(\vec{x}, t) = |\langle \vec{x} | \psi(t) \rangle|^2 = |\psi_x(\vec{x})|^2 \quad (5.89)$$

We shall see that the appropriate definition for the associated probability current is

$$j(x, t) = -\frac{i}{2} \frac{\hbar}{m} \left(\psi_x^*(x, t) \frac{\partial}{\partial x} \psi_x(x, t) - \psi_x(x, t) \frac{\partial}{\partial x} \psi_x^*(x, t) \right) \quad (5.90)$$

$$\text{or} \quad \vec{j}(\vec{x}, t) = -\frac{i}{2} \frac{\hbar}{m} \left(\psi_x^*(\vec{x}, t) \vec{\nabla} \psi_x(\vec{x}, t) - \psi_x(\vec{x}, t) \vec{\nabla} \psi_x^*(\vec{x}, t) \right) \quad (5.91)$$

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:

$$\int_V d^3x\, P(\vec{x}, t) = \int_V d^3x \langle \psi(t)\,|x\rangle\langle x\,|\psi(t)\rangle = \langle \psi(t)\,|\psi(t)\rangle \tag{5.92}$$

$$= \langle \psi(0)\,|U^\dagger(t)U(t)|\psi(0)\rangle = \langle \psi(0)\,|\psi(0)\rangle \tag{5.93}$$

Hence, if we take a time derivative of the integrated probability, we get zero: the total probability is conserved.

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.

We have

$$\frac{\partial P(\vec{x}, t)}{\partial t} = \psi_x^*(\vec{x}, t) \frac{\partial \psi_x(\vec{x}, t)}{\partial t} + \psi_x(\vec{x}, t) \frac{\partial \psi_x^*(\vec{x}, t)}{\partial t} \tag{5.94}$$

$$= \psi_x^*(\vec{x}, t) \left[-\frac{i}{\hbar} H(\vec{x}, t) \psi_x(\vec{x}, t) \right] + \psi_x(\vec{x}, t) \left[\frac{i}{\hbar} H^\dagger(\vec{x}, t) \psi_x^*(\vec{x}, t) \right] \tag{5.95}$$

$$= -\frac{i}{\hbar} \psi_x^*(\vec{x}, t) \left[-\frac{\hbar^2}{2\,m} \vec{\nabla}^2 \psi_x(\vec{x}, t) + V(\vec{x}, t) \psi_x(\vec{x}, t) \right] \tag{5.96}$$

$$+ \frac{i}{\hbar} \psi_x(\vec{x}, t) \left[-\frac{\hbar^2}{2\,m} \vec{\nabla}^2 \psi_x^*(\vec{x}, t) + V(\vec{x}, t) \psi_x^*(\vec{x}, t) \right]$$

$$= \frac{i\,\hbar}{2\,m} \left[\psi_x^*(\vec{x}, t) \vec{\nabla}^2 \psi_x(\vec{x}, t) - \psi_x(\vec{x}, t) \vec{\nabla}^2 \psi_x^*(\vec{x}, t) \right] \tag{5.97}$$

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.

Finally, pull one $\vec{\nabla}$ to the front of the expression and manipulate a bit:

$$\frac{\partial P(\vec{x}, t)}{\partial t} = \vec{\nabla} \cdot \left[\frac{i\,\hbar}{2\,m} \left[\psi_x^*(\vec{x}, t) \vec{\nabla} \psi_x(\vec{x}, t) - \psi_x(\vec{x}, t) \vec{\nabla} \psi_x^*(\vec{x}, t) \right] \right] \tag{5.98}$$
$$- \left[\frac{i\,\hbar}{2\,m} \left[\left(\vec{\nabla} \psi_x^*(\vec{x}, t) \right) \cdot \left(\vec{\nabla} \psi_x(\vec{x}, t) \right) - \left(\vec{\nabla} \psi_x(\vec{x}, t) \right) \cdot \left(\vec{\nabla} \psi_x^*(\vec{x}, t) \right) \right] \right]$$
$$= -\vec{\nabla} \cdot \vec{j}(\vec{x}, t) \tag{5.99}$$

and we have our desired result for the continuity equation.

Lecture 14:
Scattering from a Step Potential

Revision Date: 2007/10/31

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

$$V(x) = \begin{cases} 0 & x < 0 \\ V_0 & x \geq 0 \end{cases} \tag{5.100}$$

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

$$H = \frac{P^2}{2\,m} + V(x) \tag{5.101}$$

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 $V(x)$:

$$-\frac{\hbar^2}{2\,m}\frac{d^2}{dx^2}\psi_{E,x}(x) + V(x)\,\psi_{E,x}(x) = E\,\psi_{E,x}(x) \tag{5.102}$$

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:

$$\psi_{E,x,I}(x) = A \, e^{i k_1 x} + B \, e^{-i k_1 x} \qquad k_1 = \sqrt{\frac{2\,m}{\hbar^2}\,E} \tag{5.103}$$

$$\psi_{E,x,II}(x) = C \, e^{i k_2 x} + D \, e^{-i k_2 x} \qquad k_2 = \sqrt{\frac{2\,m}{\hbar^2}\,(E - V_0)} \tag{5.104}$$

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.

We note that the algebra we are about to perform would hold equally well if k_1 and/or k_2 were perfect imaginary, corresponding to a state in which the energy is less than the step height. (One keeps only k_1 and k_2 with positive imaginary part.) We will interpret these solutions after we have completed the solutions for real k_1 and k_2.

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:

$$\frac{d}{dx}\,\psi_{E,x}(x)\bigg|_{-\epsilon}^{\epsilon} = \int_{-\epsilon}^{\epsilon} dx\,\frac{d^2}{dx^2}\,\psi_{E,x}(x) \tag{5.105}$$

$$= \frac{2\,m}{\hbar^2}\int_{-\epsilon}^{0} dx\,(-E)\,\psi_{E,x,I}(x) + \frac{2\,m}{\hbar^2}\int_{0}^{\epsilon} dx\,(V_0 - E)\,\psi_{E,x,II}(x) \tag{5.106}$$

$$\approx \frac{2\,m}{\hbar^2}\,\epsilon\left[(-E)\psi_{E,x,I}(-\epsilon) + (V_0 - E)\psi_{E,x,II}(\epsilon)\right] \tag{5.107}$$

where we have used the differential 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 $\epsilon \to 0$. Thus, as we let $\epsilon \to 0$, the entire expression is guaranteed to vanish, implying continuity of $\frac{d}{dx}\,\psi_{E,x}(x)$.

So our conditions for matching the wavefunction and its derivative are:

$$A + B = C + D \qquad i\,k_1\,(A - B) = i\,k_2\,(C - D) \tag{5.108}$$

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. 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.*

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

$$1 + b_\rightarrow = c_\rightarrow \qquad k_1 (1 - b_\rightarrow) = k_2 c_\rightarrow \tag{5.109}$$

$$k_1 (1 - b_\rightarrow) = k_2 (1 + b_\rightarrow) \tag{5.110}$$

$$b_\rightarrow = \frac{k_1 - k_2}{k_1 + k_2} \qquad c_\rightarrow = \frac{2 k_1}{k_1 + k_2} \tag{5.111}$$

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.

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:

$$j_{\rightarrow,l}(x) = -\frac{i}{2}\frac{\hbar}{m}\left[\left(A^* e^{-i k_1 x} + B^* e^{i k_1 x}\right)\frac{\partial}{\partial x}\left(A e^{i k_1 x} + B e^{-i k_1 x}\right) - c.c\right] \quad (5.112)$$

$$= -\frac{i}{2}\frac{\hbar}{m}\left[\left(A^* e^{-i k_1 x} + B^* e^{i k_1 x}\right) i k_1 \left(A e^{i k_1 x} - B e^{-i k_1 x}\right)\right. \quad (5.113)$$
$$\left. - \left(A e^{i k_1 x} + B e^{-i k_1 x}\right) i k_1 \left(-A^* e^{-i k_1 x} + B^* e^{i k_1 x}\right)\right]$$

$$= \frac{1}{2}\frac{\hbar k_1}{m}\left[|A|^2 - |B|^2 + A B^* e^{2 i k_1 x} + A^* B e^{-2 i k_1 x}\right. \quad (5.114)$$
$$\left. + |A|^2 - |B|^2 - A B^* e^{2 i k_1 x} - A^* B e^{-2 i k_1 x}\right]$$

$$= \frac{\hbar k_1}{m}\left[|A|^2 - |B|^2\right] = |A|^2 \frac{\hbar k_1}{m}\left[1 - |b_\rightarrow|^2\right] \quad (5.115)$$

The same calculation for region *II* is much easier because there is only one component to the wavefunction:

$$j_{\to,II}(x) = -\frac{i}{2}\frac{\hbar}{m}\left[C^* e^{-i k_2 x}\frac{\partial}{\partial x}C e^{i k_2 x} - c.c\right] \tag{5.116}$$

$$= \frac{\hbar k_2}{m}|C|^2 = |A|^2 \frac{\hbar k_2}{m}|c_\to|^2 \tag{5.117}$$

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 position independent in each region):

$$j_{\to,I}^{in} = |A|^2 \frac{\hbar k_1}{m} \tag{5.118}$$

$$j_{\to,I}^{out} = -|A|^2 \frac{\hbar k_1}{m}|b_\to|^2 \equiv -R\, j_{\to,I}^{in} \tag{5.119}$$

$$j_{\to,II}^{out} = |A|^2 \frac{\hbar k_2}{m}|c_\to|^2 \equiv T\, j_{\to,I}^{in} \tag{5.120}$$

$$1 - R = T \tag{5.121}$$

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

$$\psi_{E,x,I}(x) = A\,e^{i\,k_1\,x} + B\,e^{-i\,k_1\,x} \qquad k_1 = \sqrt{\frac{2\,m}{\hbar^2}\,E} \tag{5.122}$$

$$\psi_{E,x,II}(x) = C\,e^{-\kappa_2\,x} + D\,e^{\kappa_2\,x} \qquad \kappa_2 = \sqrt{\frac{2\,m}{\hbar^2}\,(V_0 - E)} \tag{5.123}$$

The solution for region I is the same as Equation 5.103, while the solution for region II has changed from a complex exponential in Equation 5.104 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.

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.

Our matching conditions then become

$$A + B = C \qquad i\,k_1\,(A - B) = -\kappa_2 C \qquad (5.124)$$

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. 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:

$$1 + b_\rightarrow = c_\rightarrow \qquad i\,k_1\,(1 - b_\rightarrow) = -\kappa_2 c_\rightarrow \qquad (5.125)$$

$$i\,k_1\,(1 - b_\rightarrow) = -\kappa_2\,(1 + b_\rightarrow) \qquad (5.126)$$

$$b_\rightarrow = \frac{i\,k_1 + \kappa_2}{i\,k_1 - \kappa_2} = -\frac{(i\,k_1 + \kappa_2)^2}{k_1^2 + \kappa_2^2} = \frac{k_1^2 - \kappa_2^2}{k_1^2 + \kappa_2^2} - \frac{2\,i\,k_1\kappa_2}{k_1^2 + \kappa_2^2} \qquad (5.127)$$

$$c_\rightarrow = \frac{2\,k_1^2}{k_1^2 + \kappa_2^2} - \frac{2\,i\,k_1\kappa_2}{k_1^2 + \kappa_2^2} \qquad (5.128)$$

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.

Section 5.5 — footer

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, $\psi_{E,x,II}(x)$ and $\psi_{E,x,II}^*(x)$ differ only in their coefficents, not in the arguments of the exponentials, so

$$\psi_{E,x,II}^*(x) \frac{d}{dx} \psi_{E,x,II}(x) = \psi_{E,x,II}(x) \frac{d}{dx} \psi_{E,x,II}^*(x) \tag{5.129}$$

(the conjugation of *C* does not matter because on 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

$$j_{\rightarrow,I}^{in} = |A|^2 \frac{\hbar k_1}{m} \qquad j_{\rightarrow,I}^{out} = -|A|^2 \frac{\hbar k_1}{m} = -j_{\rightarrow,I}^{in} \qquad j_{\rightarrow,II}^{out} = 0 \tag{5.130}$$

$$R = 1 \qquad T = 0 \tag{5.131}$$

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.

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:

$$B = C \qquad \kappa_1 B = -\kappa_2 C \qquad (5.132)$$
$$\implies \kappa_1 B = -\kappa_2 B \qquad (5.133)$$

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

$$-\frac{\hbar^2}{2\,m} \frac{d^2}{dx^2} \psi_{E=0,x,I}(x) = 0 \tag{5.134}$$

This is directly integrable; the general solution is

$$\psi_{E=0,x,I}(x) = A + B\,x \tag{5.135}$$

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

$$A = C \qquad 0 = -\kappa_2 C \tag{5.136}$$

Since $\kappa_2 \neq 0$, we thus have $A = C = 0$. This ends up being the same as the $E < 0$ case.

▶ $E = V_0$: Here, the region *II* differential equation simplifies as the region *I* equation did for $E = 0$, so we have

$$\psi_{E=V_0,x,II}(x) = C \tag{5.137}$$

(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

$$A + B = C \qquad i\,k_1\,(A - B) = 0 \tag{5.138}$$

So we have $A = B$ and $C = 2\,A$. This is a rather bizarre state. The probability currents are simple and take on the same values that they do for the $0 < E < V_0$ case:

$$j^{in}_{\rightarrow,I} = |A|^2\,\frac{\hbar\,k_1}{m} \qquad j^{out}_{\rightarrow,I} = -|A|^2\,\frac{\hbar\,k_1}{m} = -j^{in}_{\rightarrow,I} \qquad j^{out}_{\rightarrow,II} = 0 \tag{5.139}$$

$$R = 1 \qquad T = 0 \tag{5.140}$$

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.

Lecture 15:
Scattering from a Step Potential: Wave Packet Example

Revision Date: 2007/11/02

Scattering from a Step Potential: Propagator

Propagator

Now that we have found the eigenstates, it is obvious to note that the unitary propagator is

$$U(t) = \int_0^\infty dk \, e^{-\frac{i}{\hbar} E_k t} \left[|\psi_{E_k}\rangle\langle\psi_{E_k}| + \theta(-k_V - k) |\psi_{E_{-k}}\rangle\langle\psi_{E_{-k}}| \right] \tag{5.141}$$

where $|\psi_{E_k}\rangle$ is the state whose position representation has wavenumber $k_1 = k$ in region *I*. 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

$$k_2 = \sqrt{\frac{2m}{\hbar^2}(E_k - V_0)} = \sqrt{k^2 - \frac{2m}{\hbar^2} V_0} \tag{5.142}$$

which picks up the appropriate factor of i when $E_k < V_0$. As for the free particle and the particle in a box, we work with the k index rather than the E index because we know how to count up the states correctly in k.

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

$$[U(t)]_{xx'} = \int_0^\infty dk\, e^{-\frac{i}{\hbar} E_k t} \left[\langle x|\psi_{E_k}\rangle\langle\psi_{E_k}|x'\rangle + \theta(-k_V - k)\langle x|\psi_{E_{-k}}\rangle\langle\psi_{E_{-k}}|x'\rangle \right]$$
(5.143)

$$= \int_0^\infty dk\, e^{-\frac{i}{\hbar} E_k t} \left[\psi_{E_k,x}^*(x)\,\psi_{E_k,x}(x') + \theta(-k_V - k)\,\psi_{E_{-k},x}^*(x)\,\psi_{E_{-k},x}(x') \right]$$
(5.144)

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 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.

Scattering from a Step Potential: Wave-Packet Example

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 try to 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

$$\langle x \,|\psi(0)\, \rangle = \left(\frac{1}{2 \, \pi \, \sigma_x^2} \right)^{1/4} e^{\frac{i}{\hbar} \, p_0 \, (x+a)} e^{-\frac{(x+a)^2}{4 \, \sigma_x^2}} \equiv \psi_i(x, 0) \tag{5.145}$$

with expectation values at $t = 0$ of

$$\langle X \rangle = -a \qquad \langle (\Delta X)^2 \rangle = \sigma_x^2 \qquad \langle P \rangle = p_0 = \hbar \, k_0 \qquad \langle (\Delta P)^2 \rangle = \sigma_p^2 = \frac{\hbar^2}{4 \, \sigma_x^2} \tag{5.146}$$

Recall that our σ_x is related to Shankar's Δ_x by $\Delta_x = \sigma_x \sqrt{2}$.

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 $+\infty$ that we will call ψ_t and one that goes off to $-\infty$ that we will call ψ_r. We will calculate the probability of transmission and reflection by calculate 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

$$\sigma_x \ll a \tag{5.147}$$

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

$$\sigma_p \ll p_0 \tag{5.148}$$

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 5.30, 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

$$\frac{a}{p_0/m} \ll \frac{\sigma_x}{\sigma_p/m} \tag{5.149}$$

$$\frac{\sigma_p}{p_0} \ll \frac{\sigma_x}{a} \tag{5.150}$$

So, if we define

$$\alpha = \frac{\sigma_p}{p_0} \qquad \beta = \frac{\sigma_x}{a} \tag{5.151}$$

then we require

$$\alpha \ll 1 \qquad \beta \ll 1 \qquad \alpha \ll \beta \tag{5.152}$$

which is entirely feasible.

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:

$$\langle E \rangle = \frac{1}{2\,m} \langle P^2 \rangle = \frac{1}{2\,m} \left(\langle P \rangle^2 + \langle (\Delta P)^2 \rangle \right) = \frac{1}{2\,m} \left(p_0^2 + \sigma_p^2 \right) \tag{5.153}$$

$$\sigma_E^2 = \langle (\Delta E)^2 \rangle = \left(\frac{1}{2\,m} \right)^2 \left(\langle P^4 \rangle - \langle P^2 \rangle^2 \right) \tag{5.154}$$

$$= \left(\frac{1}{2\,m} \right)^2 \left(\langle P \rangle^4 + 6\langle P \rangle^2 \langle (\Delta P)^2 \rangle + 3\langle (\Delta P)^2 \rangle^2 - \left(\langle P \rangle^2 + \langle (\Delta P)^2 \rangle \right)^2 \right) \tag{5.155}$$

$$= \left(\frac{1}{2\,m} \right)^2 \left(4\langle P \rangle^2 \langle (\Delta P)^2 \rangle + 2\langle (\Delta P)^2 \rangle^2 \right) = \left(\frac{1}{2\,m} \right)^2 \left(4\,p_0^2\,\sigma_p^2 + 2\,\sigma_p^4 \right) \tag{5.156}$$

$$\frac{\sigma_E}{E} = \frac{\sigma_p \sqrt{4\,p_0^2 + 2\,\sigma_p^2}}{p_0^2 + \sigma_p^2} \approx 2\,\frac{\sigma_p}{p_0} \ll 1 \tag{5.157}$$

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 $\sigma_p \ll p_0$.

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 5.141 and 5.144). The first step in applying the propagator is, as always, to take the matrix elements $\langle \psi_{E_k} | \psi_i(0) \rangle$. 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

$$\langle \psi_{E_k} | \psi_i(0) \rangle = \int_{-\infty}^{\infty} dx \, \langle \psi_{E_k} | x \rangle \langle x | \psi_i(0) \rangle \tag{5.158}$$

$$= \int_{-\infty}^{\infty} dx \, \psi_{E_k, x}^*(x) \, \psi_i(x, 0) \tag{5.159}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx \, \Big[\left(e^{ikx} + b_\rightarrow(k) \, e^{-ikx} \right) \theta(-x) \tag{5.160}$$
$$+ c_\rightarrow(k) \, e^{ik_2(k)x} \theta(x) \Big]^* \psi_i(x, 0)$$

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. The normalization of $\psi_{E_k, x}(x)$ is set by analogy to free-particle states, which need a $1/\sqrt{2\pi}$ normalization factor when normalized to a delta function. The θ functions let us write our region I and II solutions together in concise form.

Now, we may make some simplifications. By construction, the initial wavefunction $\psi_i(x, 0)$ is essentially zero in region *II* ($x > 0$) where $\theta(x) \neq 0$; so the integral with the $c_\rightarrow$ will vanish. By the same argument, we may drop $\theta(-x)$ as superfluous. Next, let's rewrite the above as follows:

$$\langle \psi_{E_k} | \psi_i(0) \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\, e^{-ikx} \psi_i(x, 0) \tag{5.161}$$

$$+ \frac{b_\rightarrow^*(k)}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\, e^{ikx} \psi_i(x, 0) \tag{5.162}$$

The two terms are calculating the Fourier transform at k of $\psi_i(x, 0)$ 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.)

We know from our discussion of Gaussian wave packets for the free particle that the Fourier transform of our initial state is

$$\langle k \,|\psi_i(0)\,\rangle = \frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{\infty} dx \, e^{-i\,k\,x} \left(\frac{1}{2\,\pi\,\sigma_x^2}\right)^{1/4} e^{i\,k_0\,(x+a)} e^{-\frac{(x+a)^2}{4\,\sigma_x^2}} \tag{5.163}$$

$$= \left(\frac{1}{2\,\pi\,\sigma_k^2}\right)^{1/4} e^{-\frac{(k-k_0)^2}{4\,\sigma_k^2}} e^{i\,k\,a} \tag{5.164}$$

where we have made the simple change of variables to $k = p/\hbar$ to match up with the way we index the states here k. Since $\sigma_k/k_0 = \sigma_p/p_0 \ll 1$, it holds that the Gaussian essentially vanishes for $k < 0$: our initial contains essentially no components in the initially left-going states.

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)

$$\langle \psi_{E_{-k}} \,|\, \psi_i(0) \,\rangle = \int_{-\infty}^{\infty} dx \, \psi_{E_{-k}, x}^*(x) \, \psi_i(x, 0) \tag{5.165}$$

$$= \frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{\infty} dx \left[\left(e^{-i\,k_2(k)\,x} + c_\leftarrow(k)\, e^{i\,k_2(k)\,x} \right) \theta(x) \right. \tag{5.166}$$
$$\left. + b_\leftarrow(k)\, e^{-i\,k\,x}\, \theta(-x) \right]^* \psi_i(x, 0)$$

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 zero Fourier transform for $k < 0$, 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.

We have our inner products, so now the state at some time t later (for any $E > 0$) is

$$\langle x \,|\psi_i(t)\,\rangle = \int_0^\infty dk \, e^{-\frac{i}{\hbar} E_k t} \langle x \,|\psi_{E_k}\,\rangle \langle \psi_{E_k} \,|\psi_i(0)\,\rangle \qquad (5.167)$$

$$= \int_0^\infty dk \, e^{-\frac{i}{\hbar} E_k t} \frac{1}{\sqrt{2\pi}} \Big[\left(e^{ikx} + b_\rightarrow(k)\, e^{-ikx} \right) \theta(-x) \qquad (5.168)$$
$$+ c_\rightarrow(k)\, e^{ik_2(k)x}\, \theta(x) \Big]$$
$$\left(\frac{1}{2\pi\sigma_k^2} \right)^{1/4} e^{-\frac{(k-k_0)^2}{4\sigma_k^2}} e^{ika}$$

For $0 < E \leq V_0$, the last term goes from a complex exponential to a real exponential as necessary.

For $E > 0$, we can convert the above integral, with some work, to a set of reverse Fourier transforms of the Gaussian.

We first 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 three integrals:

$$\langle x \,|\psi_i(t)\,\rangle = \frac{\theta(-x)}{\sqrt{2\pi}} \int_0^\infty dk\, e^{-\frac{i}{\hbar} E_k t} e^{i k x} \left(\frac{1}{2\pi\sigma_k^2}\right)^{1/4} e^{-\frac{(k-k_0)^2}{4\sigma_k^2}} e^{i k a} \qquad (5.169)$$

$$+ \frac{\theta(-x)\, b_\rightarrow(k_0)}{\sqrt{2\pi}} \int_0^\infty dk\, e^{-\frac{i}{\hbar} E_k t} e^{-i k x} \left(\frac{1}{2\pi\sigma_k^2}\right)^{1/4} e^{-\frac{(k-k_0)^2}{4\sigma_k^2}} e^{i k a}$$

$$+ \frac{\theta(x)\, c_\rightarrow(k_0)}{\sqrt{2\pi}} \int_0^\infty dk\, e^{-\frac{i}{\hbar} E_k t} e^{i k_2(k) x} \left(\frac{1}{2\pi\sigma_k^2}\right)^{1/4} e^{-\frac{(k-k_0)^2}{4\sigma_k^2}} e^{i k a}$$

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 $\langle x | \psi_i(t) \rangle$. 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.

Section 5.5   Simple One-Dimensional Problems: Scattering from a Step Potential   Page 264

To evaluate the second term, let's change the integration variable from k to $k' = -k$:

$$\text{2nd term} = \frac{\theta(-x)\, b_{\rightarrow}(k_0)}{\sqrt{2\,\pi}} \int_{-\infty}^{0} dk'\, e^{-\frac{i}{\hbar} E_k t} e^{i k' x} \left(\frac{1}{2\,\pi\,\sigma_k^2}\right)^{1/4} e^{-\frac{(k' + k_0)^2}{4\,\sigma_k^2}} e^{-i k' a}$$

(5.170)

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 a at $t = 0$. 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, mean momentum $p_0 = \hbar k_0$, and momentum spread $\sigma_p = \hbar \sigma_k << p_0$. The $\theta(-x)$ function now 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

$$\text{2nd term} = \left[\frac{1}{2\,\pi\,\sigma_x^2 \left(1 + \frac{i\,\hbar\,t}{2\,m\,\sigma_x^2}\right)}\right]^{-1/4} \exp\left[-\frac{\left(x + \frac{\hbar\,k_0}{m}\,t\right)^2}{4\,\sigma_x^2 \left(1 + \frac{i\,\hbar\,t}{2\,m\,\sigma_x^2}\right)}\right]$$

$$\exp\left(-i\,k_0\,x\right) \exp\left(-\frac{i}{\hbar}\,E_{k_0}\,t\right)$$

(5.171)

The third term is much harder to evaluate. But we don't need to. We just want to know 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 just

$$R = |b_{\rightarrow}(k_0)|^2 \tag{5.172}$$

By conservation of probability, we know that the probability in the transmitted wave is

$$T = 1 - R \tag{5.173}$$

We know that $|b_{\rightarrow}(k_0)| = 1$ for $0 < E < V_0$ by our direct calculation of $b_{\rightarrow}(k)$ earlier (Equations 5.111, 5.127, and 5.128), 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 16:
Theorems on One-Dimensional States
Simple Harmonic Oscillator: Coordinate Basis

Revision Date: 2007/11/05

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

$$\langle x|H|\psi_{E,1}\rangle = \langle x|E|\psi_{E,1}\rangle \qquad \langle x|H|\psi_{E,2}\rangle = \langle x|E|\psi_{E,2}\rangle \qquad (5.174)$$

We may write the above as (recall, use the $\{|x\rangle\}$-basis matrix elements of H and insert completeness in this basis):

$$-\frac{\hbar^2}{2\,m}\frac{d^2}{dx^2}\psi_{E,1,x}(x) + V(x)\,\psi_{E,1,x}(x) = E\,\psi_{E,1,x}(x) \qquad (5.175)$$

$$-\frac{\hbar^2}{2\,m}\frac{d^2}{dx^2}\psi_{E,2,x}(x) + V(x)\,\psi_{E,2,x}(x) = E\,\psi_{E,2,x}(x) \qquad (5.176)$$

Multiply the first equation by $\psi_{E,2,x}(x)$ and the second by $\psi_{E,1,x}(x)$ and difference to obtain

$$\psi_{E,1,x}(x)\,\frac{d^2}{dx^2}\,\psi_{E,2,x}(x) - \psi_{E,2,x}(x)\,\frac{d^2}{dx^2}\,\psi_{E,1,x}(x) = 0 \tag{5.177}$$

$$\frac{d}{dx}\left(\psi_{E,1,x}(x)\,\frac{d}{dx}\,\psi_{E,2,x}(x) - \psi_{E,2,x}(x)\,\frac{d}{dx}\,\psi_{E,1,x}(x)\right) = 0 \tag{5.178}$$

$$\psi_{E,1,x}(x)\,\frac{d}{dx}\,\psi_{E,2,x}(x) - \psi_{E,2,x}(x)\,\frac{d}{dx}\,\psi_{E,1,x}(x) = c \tag{5.179}$$

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.

With $c = 0$, we have

$$\frac{1}{\psi_{E,1,x}(x)} \frac{d}{dx} \psi_{E,1,x}(x) = \frac{1}{\psi_{E,2,x}(x)} \frac{d}{dx} \psi_{E,2,x}(x) \tag{5.180}$$

$$\frac{d}{dx} \left[\ln \psi_{E,1,x}(x) \right] = \frac{d}{dx} \left[\ln \psi_{E,2,x}(x) \right] \tag{5.181}$$

$$\ln \psi_{E,1,x}(x) = \ln \psi_{E,2,x}(x) + d \tag{5.182}$$

$$\psi_{E,1,x}(x) = e^d \, \psi_{E,2,x} \tag{5.183}$$

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.

The eigenfunctions of $H = \frac{P^2}{2m} + V(X)$ 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:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi_{E,x}(x) + V(x) \, \psi_{E,x}(x) = E \, \psi_{E,x}(x) \qquad (5.184)$$

For Hamiltonians of the above form, we can complex conjugate the above equation:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi_{E,x}^*(x) + V(x) \, \psi_{E,x}^*(x) = E \, \psi_{E,x}^*(x) \qquad (5.185)$$

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.

The fact that H stayed the same under complex conjugation is a result of its particular form; if one had $H = \frac{P^2}{2m} - \beta \frac{P}{m} + V(X)$ (*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:

$$\psi_{E,x,r}(x) = \frac{1}{2}\left(\psi_{E,x}(x) + \psi_{E,x}^*(x)\right) \qquad \psi_{E,x,i}(x) = \frac{1}{2i}\left(\psi_{E,x}(x) - \psi_{E,x}^*(x)\right) \tag{5.186}$$

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 $\psi_{E,x}^*(x)$.

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

$$\psi_{E,x,r}(x) = \frac{1}{2}\left(1 + e^{i\theta}\right)\psi_{E,x}(x) \qquad \psi_{E,x,i}(x) = \frac{1}{2i}\left(1 - e^{i\theta}\right)\psi_{E,x}(x) \quad (5.187)$$

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

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

$$V(x) = V(x_0) + \frac{d}{dx} V(x)\Big|_{x_0} (x - x_0) + \frac{1}{2} \frac{d^2}{dx^2} V(x)\Big|_{x_0} (x - x_0)^2 + \cdots \qquad (6.1)$$

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 $\frac{p^2}{2m}$, then the Hamiltonian for the system may be approximated as

$$\mathcal{H}(x, p) = \frac{p^2}{2m} + \frac{1}{2} k x^2 \qquad (6.2)$$

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.

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

$$\mathcal{H}(x, p) = \frac{p^2}{2\,m} + \frac{1}{2}\,m\,\omega^2\,x^2 \qquad \Longrightarrow \qquad H(X, P) = \frac{P^2}{2\,m} + \frac{1}{2}\,m\,\omega^2\,X^2 \qquad (6.3)$$

As usual, we first need to find the eigenstates of the Hamiltonian, $H|\psi_E\rangle = E|\psi_E\rangle$.
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!):

$$\langle x|H|\psi_E\rangle = \langle x|E|\psi_E\rangle \qquad (6.4)$$

$$\int_{-\infty}^{\infty} dx'\,\langle x|H|x'\rangle\langle x'|\psi_E\rangle = E\,\langle x|\psi_E\rangle \qquad (6.5)$$

$$\int_{-\infty}^{\infty} dx'\,\langle x|\left(\frac{P^2}{2\,m} + \frac{1}{2}\,m\,\omega^2\,X^2\right)|x'\rangle\,\psi_{E,x}(x') = E\,\psi_{E,x}(x) \qquad (6.6)$$

We calculated in Equations 5.38 and 5.48 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:

$$\int_{-\infty}^{\infty} dx\, \delta(x - x') \left[-\frac{\hbar^2}{2m} \frac{d}{d(x')^2} + \frac{1}{2} m\omega^2 (x')^2 \right] \psi_{E,x}(x') = E\, \psi_{E,x}(x) \quad (6.7)$$

$$-\frac{\hbar^2}{2m} \frac{d}{dx^2} \psi_{E,x}(x) + \frac{1}{2} m\omega^2 x^2 \psi_{E,x}(x) = E\, \psi_{E,x}(x) \quad (6.8)$$

There are two natural scales in the problem, 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:

$$b^2 \frac{d}{dx^2} \psi_{E,x}(x) + \frac{2mEb^2}{\hbar^2} \psi_{E,x}(x) - \frac{m^2\omega^2 b^4}{\hbar^2} \frac{x^2}{b^2} \psi_{E,x}(x) = 0 \quad (6.9)$$

$$\frac{d^2}{dy^2} \psi_{\varepsilon}(y) + \left(2\varepsilon - y^2 \right) \psi_{\varepsilon}(y) = 0 \quad (6.10)$$

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.

Solving the Differential Equation

We must now find the solutions of

$$\frac{d^2}{dy^2}\,\psi_\varepsilon(y) + \left(2\,\varepsilon - y^2\right)\psi_\varepsilon(y) = 0 \tag{6.11}$$

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 good to get good 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

$$\frac{d^2}{dy^2} \psi_\varepsilon(y) - y^2 \psi_\varepsilon(y) = 0 \tag{6.12}$$

(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

$$\psi(y) = A y^m e^{\pm \frac{y^2}{2}} \tag{6.13}$$

which we can see by direct substitution:

$$\frac{d^2}{dy^2} \psi(y) = \frac{d}{dy} \left(\left[m y^{m-1} + y^m (\pm y) \right] A e^{\pm \frac{y^2}{2}} \right) \tag{6.14}$$

$$= \left[\left(m(m-1) y^{m-2} \pm (m+1) y^m \right) + \left(m y^{m-1} \pm y^{m+1} \right) (\pm y) \right] A e^{\pm \frac{y^2}{2}}$$

$$= \left[y^{m+2} \pm (2m+1) y^m + m(m-1) y^{m-2} \right] A e^{\pm \frac{y^2}{2}}$$

Now take the $|y| \to \infty$ limit:

$$\frac{d^2}{dy^2} \psi(y) \stackrel{|y| \to \infty}{\longrightarrow} y^2 A y^m e^{\pm \frac{y^2}{2}} = y^2 \psi(y)$$

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 $y^m e^{-\frac{y^2}{2}}$.

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

$$\frac{d^2}{dy^2} \psi_\varepsilon(y) + 2\,\varepsilon\,\psi_\varepsilon(y) = 0 \tag{6.15}$$

This is a second-order linear differential equation with constant coefficients, so we know the solution is a sum of harmonic functions:

$$\psi_\varepsilon(y) = \alpha \cos\left(\sqrt{2\,\varepsilon}\,y\right) + \beta \sin\left(\sqrt{2\,\varepsilon}\,y\right) \tag{6.16}$$

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

$$\psi_\varepsilon(y) \stackrel{|y| \to 0}{\longrightarrow} \alpha + \beta\,\sqrt{2\,\varepsilon}\,y \tag{6.17}$$

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

$$\psi_\varepsilon(y) = u_\varepsilon(y)\, e^{-\frac{y^2}{2}} \tag{6.18}$$

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 $u(y)$:

$$\left[\frac{d^2}{dy^2} + \left(2\,\varepsilon - y^2\right)\right]\left(u_\varepsilon(y)\, e^{-\frac{y^2}{2}}\right) = 0 \tag{6.19}$$

$$\left[\frac{d^2}{dy^2}\, u_\varepsilon(y) - 2\, y\, \frac{d}{dy}\, u_\varepsilon(y) + \left(y^2 - 1 + 2\,\varepsilon - y^2\right) u_\varepsilon(y)\right] e^{-\frac{y^2}{2}} = 0 \tag{6.20}$$

$$\frac{d^2}{dy^2}\, u_\varepsilon(y) - 2\, y\, \frac{d}{dy}\, u_\varepsilon(y) + \left(2\,\varepsilon - 1\right) u_\varepsilon(y) = 0 \tag{6.21}$$

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

$$\sum_{n=0}^{\infty} C_{\varepsilon,n} \left[n(n-1) y^{n-2} - 2n y^n + (2\varepsilon - 1) y^n \right] = 0 \tag{6.22}$$

$$\sum_{m=0}^{\infty} C_{\varepsilon,m+2} (m+2)(m+1) y^m = \sum_{n=0}^{\infty} C_{\varepsilon,n} (2n + 1 - 2\varepsilon) y^n \tag{6.23}$$

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

$$C_{\varepsilon,n+2} = C_{\varepsilon,n} \frac{2n + 1 - 2\varepsilon}{(n+2)(n+1)} \tag{6.24}$$

The coefficients $C_{\varepsilon,0}$ and $C_{\varepsilon,1}$ are left to be determined by initial conditions.

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 $\sum_{n=0}^{\infty} C_n y^n \stackrel{|y| \to \infty}{\longrightarrow} y^m$ 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:

$$0 = C_{\varepsilon,n+2} = C_{\varepsilon,n} \frac{2n+1-2\varepsilon}{(n+2)(n+1)} \tag{6.25}$$

$$0 = 2n+1-2\varepsilon \tag{6.26}$$

$$\varepsilon = \frac{2n+1}{2} \tag{6.27}$$

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.

Lecture 17:
Simple Harmonic Oscillator:
Coordinate Basis Continued
Energy Basis

Revision Date: 2007/11/07

Coordinate Basis

Explicit Form for the SHO Solutions

Our solutions are of the form

$$\psi_{E_n,x}(x) = \left(\frac{m\,\omega}{\pi\,\hbar\,2^{2n}\,(n!)^2}\right)^{1/4} H_n\left(x\sqrt{\frac{m\,\omega}{\hbar}}\right) e^{-\frac{m\,\omega\,x^2}{2\,\hbar}} \qquad E_n = \left(n + \frac{1}{2}\right)\hbar\omega \tag{6.28}$$

where $H_n(y)$ are the Hermite polynomials

$$H_0(y) = 1 \qquad\qquad H_1(y) = 2\,y \tag{6.29}$$

$$H_2(y) = -2\left(1 - 2\,y^2\right) \qquad H_3(y) = -12\left(y - \frac{2}{3}\,y^3\right) \tag{6.30}$$

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.

Coordinate Basis (cont.)

The Propagator

We calculate the propagator by the standard formula, Equation 4.32, and also calculate its $\{|x\rangle\}$-basis matrix elements are:

$$U(t) = e^{-\frac{i}{\hbar} H t} = \sum_n e^{-\frac{i}{\hbar} E_n t} |\psi_{E_n}(0)\rangle\langle\psi_{E_n}(0)| = \sum_{n=0}^{\infty} e^{-i\left(n+\frac{1}{2}\right)\omega t} |\psi_{E_n}\rangle\langle\psi_{E_n}| \tag{6.31}$$

$$[U(t)]_{xx'} = \langle x|U(t)|x'\rangle = \sum_{n=0}^{\infty} e^{-i\left(n+\frac{1}{2}\right)\omega t}\langle x|\psi_{E_n}\rangle\langle\psi_{E_n}|x'\rangle \tag{6.32}$$

$$= \sum_{n=0}^{\infty} e^{-i\left(n+\frac{1}{2}\right)\omega t}\left(\frac{m\omega}{\pi\hbar 2^{2n}(n!)^2}\right)^{1/2} H_n(x)\, e^{-\frac{m\omega x^2}{2\hbar}} H_n(x')\, e^{-\frac{m\omega (x')^2}{2\hbar}} \tag{6.33}$$

$$= \left(\frac{m\omega}{2\pi i\hbar\sin\omega t}\right)^{1/2} \exp\left[\frac{i}{\hbar} m\omega \frac{(x^2+(x')^2)\cos\omega t - 2xx'}{2\sin\omega t}\right] \tag{6.34}$$

where the sum is not obvious. Like Shankar, we will use path integrals to do it later.

Properties of the SHO Solutions

We note some interesting properties of the solutions:

▶ **Parity**
The parity operator Π makes the transformation on the wavefunction $x \to -x$.
We will discuss it more detail later (in particular, its formal form). One thing we
notice is that the symmetry of the potential $V(x)$ implies that the Hamiltonian and
the parity operator commute, so the eigenfunctions of H must also be
eigenfunctions of Π. Since $\Pi^2 = I$, 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

$$\frac{d}{dy}\,\psi_\varepsilon(y')\bigg|_y^\infty = \int_y^\infty \frac{d^2}{d(y')^2}\,\psi_\varepsilon(y') = \int_y^\infty \left((y')^2 - 2\,\varepsilon\right)\psi_\varepsilon(y') \quad (6.35)$$

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

$$\frac{d}{dy}\,\psi_\varepsilon(y) = -\int_y^\infty \left[(y')^2 - 2\,\varepsilon\right]\psi_\varepsilon(y') \quad (6.36)$$

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 $\psi_\varepsilon(y)$. 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

$$x_n < y_0\sqrt{\frac{\hbar}{m\,\omega}} = \sqrt{\frac{\hbar}{m\,\omega}}\sqrt{2\left(n + \frac{1}{2}\right)} = \sqrt{\frac{2\,E_n}{k}} = x_0 \qquad (6.37)$$

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.

▶ **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 n modes do not satisfy this.

Energy Basis

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

$$H = \sum_n E_n |\psi_{E_n}\rangle\langle\psi_{E_n}| \equiv \sum_n E_n |E_n\rangle\langle E_n| \equiv \sum_n E_n |n\rangle\langle n| \qquad (6.38)$$

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

$$a = \sqrt{\frac{1}{2}\,\frac{m\,\omega}{\hbar}}\,X + i\sqrt{\frac{1}{2}\,\frac{1}{m\,\omega\,\hbar}}\,P \qquad (6.39)$$

Its adjoint is

$$a^\dagger = \sqrt{\frac{1}{2}\,\frac{m\,\omega}{\hbar}}\,X - i\sqrt{\frac{1}{2}\,\frac{1}{m\,\omega\,\hbar}}\,P \qquad (6.40)$$

(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

$$\left[a, a^\dagger\right] = 1 \qquad H = \left(a^\dagger a + \frac{1}{2}\right)\hbar\,\omega \quad \Longleftrightarrow \quad \widehat{H} = \frac{H}{\hbar\,\omega} = \left(a^\dagger a + \frac{1}{2}\right) \qquad (6.41)$$

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.

Energy Basis (cont.)

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

$$\widehat{H}|\varepsilon\rangle = \varepsilon|\varepsilon\rangle \tag{6.42}$$

(The use of the ε symbol is suggestive.) First, we need the following:

$$[a, \widehat{H}] = \left[a, a^{\dagger}a + \frac{1}{2}\right] = aa^{\dagger}a - a^{\dagger}aa = \left([a, a^{\dagger}] + a^{\dagger}a\right)a - a^{\dagger}aa = a \tag{6.43}$$

$$[a^{\dagger}, \widehat{H}] = \left[a^{\dagger}, a^{\dagger}a + \frac{1}{2}\right] = a^{\dagger}a^{\dagger}a - a^{\dagger}aa^{\dagger} = a^{\dagger}a^{\dagger}a - a^{\dagger}\left([a, a^{\dagger}] + a^{\dagger}a\right) = -a^{\dagger} \tag{6.44}$$

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 $\widehat{H}$:

$$\widehat{H}\,a\,|\varepsilon\rangle = \left([\widehat{H},a] + a\,\widehat{H}\right)|\varepsilon\rangle = (-a + a\,\varepsilon)\,|\varepsilon\rangle = (\varepsilon - 1)\,a\,|\varepsilon\rangle \tag{6.45}$$

$$\widehat{H}\,a^\dagger|\varepsilon\rangle = \left([\widehat{H},a^\dagger] + a^\dagger\,\widehat{H}\right)|\varepsilon\rangle = \left(a^\dagger + a^\dagger\,\varepsilon\right)|\varepsilon\rangle = (\varepsilon + 1)\,a^\dagger|\varepsilon\rangle \tag{6.46}$$

That is

$$a\,|\varepsilon\rangle = C_\varepsilon|\varepsilon - 1\rangle \qquad a^\dagger|\varepsilon\rangle = C_{\varepsilon+1}|\varepsilon + 1\rangle \tag{6.47}$$

(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 $|\psi\rangle$:

$$\langle\psi|H|\psi\rangle = \frac{1}{2\,m}\langle\psi|P^2|\psi\rangle + \frac{1}{2}\,k\langle\psi|X^2|\psi\rangle = \frac{1}{2\,m}|P|\psi\rangle|^2 + \frac{1}{2}\,k|X|\psi\rangle|^2 \geq 0 \tag{6.48}$$

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

$$a|\varepsilon_0\rangle = |0\rangle \tag{6.49}$$

so that we cannot obtain any lower energy states. We then have

$$\langle\varepsilon_0|a^\dagger a|\varepsilon_0\rangle = \langle 0|0\rangle = 0 \implies \langle\varepsilon_0|\widehat{H}|\varepsilon_0\rangle = \frac{1}{2} \text{ and } \langle\varepsilon_0|H|\varepsilon_0\rangle = \frac{1}{2}\,\hbar\omega \tag{6.50}$$

where we have used $\langle 0|0\rangle = 0$ according to the standard rules for inner products.

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.

Energy Basis (cont.)

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 vector 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 consider 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.

We have not formally defined the inner product on this space. 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 can are normalized. Since the C_ε constants are still undetermined, we still have the 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:

$$1 = \langle\varepsilon - 1\,|\varepsilon - 1\rangle = |C_\varepsilon|^{-2}\langle\varepsilon\,|a^\dagger a\,|\varepsilon\rangle = |C_\varepsilon|^{-2}\langle\varepsilon\,|\left(\widehat{H} - \frac{1}{2}\right)|\varepsilon\rangle \tag{6.51}$$

$$= |C_\varepsilon|^{-2}\left(\varepsilon - \frac{1}{2}\right)\langle\varepsilon\,|\varepsilon\rangle = |C_\varepsilon|^{-2}\left(\varepsilon - \frac{1}{2}\right) \tag{6.52}$$

$$\implies \quad |C_\varepsilon|^2 = \varepsilon - \frac{1}{2} \tag{6.53}$$

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

$$|C_n|^2 = n \qquad \Longleftrightarrow \qquad C_n = e^{i\,\phi_n}\sqrt{n} \tag{6.54}$$

where ϕ_n is arbitrary for each n. The simplest convention is $\phi_n = 0$ for all n, giving

$$a\,|n\rangle = \sqrt{n}\,|n-1\rangle \qquad a^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle \qquad \langle n|m\rangle = \delta_{nm} \tag{6.55}$$

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:

$$|n\rangle = \frac{1}{\sqrt{n!}}\left(a^\dagger\right)^n|0\rangle \tag{6.56}$$

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* $|0\rangle$. 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:

$$\langle n|a|m\rangle = \sqrt{m}\langle n|m-1\rangle = \sqrt{m}\,\delta_{n,m-1} \tag{6.57}$$

$$\langle n|a^\dagger|m\rangle = \sqrt{m+1}\langle n|m+1\rangle = \sqrt{m+1}\,\delta_{n,m+1} \tag{6.58}$$

$$X = \sqrt{\frac{\hbar}{2\,m\,\omega}}\left(a^\dagger + a\right) \implies \langle n|X|m\rangle = \sqrt{\frac{\hbar}{2\,m\,\omega}}\left(\sqrt{m+1}\,\delta_{n,m+1} + \sqrt{m}\,\delta_{n,m-1}\right) \tag{6.59}$$

$$P = i\sqrt{\frac{m\,\omega\,\hbar}{2}}\left(a^\dagger - a\right) \implies \langle n|P|m\rangle = i\sqrt{\frac{m\,\omega\,\hbar}{2}}\left(\sqrt{m+1}\,\delta_{n,m+1} - \sqrt{m}\,\delta_{n,m-1}\right) \tag{6.60}$$

$$H = \left(N + \frac{1}{2}\right)\hbar\,\omega \implies \langle n|H|m\rangle = \left(n + \frac{1}{2}\right)\hbar\,\omega\,\delta_{n,m} \tag{6.61}$$

The marix representations are given in Shankar; they are obvious from the above.

Shankar makes the point that working in the energy basis, which 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 evaluation 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 18:
Simple Harmonic Oscillator:
Energy Basis – Coordinate Basis Correspondence
Rewriting Postulate 2

Revision Date: 2007/11/09

Energy Basis – Coordinate Basis Correspondence

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:

$$0 = \langle x|a|0\rangle = \int_{-\infty}^{\infty} dx'\, \langle x|a|x'\rangle\langle x'|0\rangle \tag{6.62}$$

$$= \int_{-\infty}^{\infty} dx'\, \langle x|\left(\sqrt{\frac{m\omega}{2\hbar}}\,X + i\sqrt{\frac{1}{2m\omega\hbar}}\,P\right)|x'\rangle\,\psi_{0,x}(x') \tag{6.63}$$

$$= \sqrt{\frac{m\omega}{2\hbar}}\int_{-\infty}^{\infty} dx'\,\left(x\,\delta(x-x') + \frac{\hbar}{m\omega}\left[\frac{d}{dx}\delta(x-x')\right]\right)\psi_{0,x}(x') \tag{6.64}$$

$$= \sqrt{\frac{m\omega}{2\hbar}}\int_{-\infty}^{\infty} dx'\,\left(x\,\delta(x-x') + \frac{\hbar}{m\omega}\,\delta(x-x')\frac{d}{dx'}\right)\psi_{0,x}(x') \tag{6.65}$$

$$= \sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{\hbar}{m\omega}\frac{d}{dx}\right)\psi_{0,x}(x) \tag{6.66}$$

We thus have a first-order differential equation for $\psi_{0,x}(x)$:

$$\frac{1}{\sqrt{2}}\left(\sqrt{\frac{m\omega}{\hbar}}\,x + \sqrt{\frac{\hbar}{m\omega}}\,\frac{d}{dx}\right)\psi_{0,x}(x) = 0 \iff \frac{1}{\sqrt{2}}\left(y + \frac{d}{dy}\right)\psi_0(y) = 0 \quad (6.67)$$

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:

$$\frac{d\psi_0(y)}{\psi_0(y)} = -y\,dy \tag{6.68}$$

$$\ln\psi_0(y) + c = -\frac{y^2}{2} \tag{6.69}$$

$$\psi_0(y) = A\,e^{-\frac{y^2}{2}} \tag{6.70}$$

We thus recover the $\{|x\rangle\}$-basis representation of the ground state!

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:

$$\langle x | n \rangle = \langle x | \frac{\left(a^{\dagger}\right)^{n}}{\sqrt{n!}} | 0 \rangle \tag{6.71}$$

$$= \frac{1}{\sqrt{n!}} \left(\sqrt{\frac{m\omega}{2\hbar}} X - i\sqrt{\frac{1}{2m\omega\hbar}} P \right)^{n} \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} \tag{6.72}$$

$$= \frac{1}{\sqrt{n!}} \left(\sqrt{\frac{m\omega}{2\hbar}} x - \hbar\sqrt{\frac{1}{2m\omega\hbar}} \frac{d}{dx} \right)^{n} \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} \tag{6.73}$$

$$= \left(\frac{m\omega}{\pi\hbar 2^{2n}(n!)^2} \right)^{1/4} \left(y - \frac{d}{dy} \right)^{n} e^{-\frac{y^2}{2}} \tag{6.74}$$

(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 6.28) by using recursion relations for the Hermite polynomials. The two recursion relations given in Shankar are

$$\frac{d}{dy} H_n(y) = 2\,n\,H_{n-1}(y) \qquad H_{n+1}(y) = 2\,y\,H_n(y) - 2\,n\,H_{n-1}(y) \qquad (6.75)$$

which can clearly be combined to find

$$H_{n+1}(y) = 2\,y\,H_n(y) - \frac{d}{dy} H_n(y) \qquad (6.76)$$

Let's do the proof of equivalence inductively. First, show it is true for $n = 1$:

$$\langle x\,|1\,\rangle = \left(\frac{m\,\omega}{\pi\,\hbar\,2^2\,(1\,!)^2}\right)^{1/4} \left(y - \frac{d}{dy}\right) e^{-\frac{y^2}{2}} = \left(\frac{m\,\omega}{\pi\,\hbar\,2^2\,(1\,!)^2}\right)^{1/4} (2y)\,e^{-\frac{y^2}{2}} \qquad (6.77)$$

$$= \left(\frac{m\,\omega}{\pi\,\hbar\,2^2\,(1\,!)^2}\right)^{1/4} H_1(y)\,e^{-\frac{y^2}{2}} = \psi_{E_1,x}(y) \qquad (6.78)$$

as expected.

Next, assume it is true for *n* and show that it is true for $n+1$:

$$\langle x \,|\, n+1 \rangle = \left(\frac{m\omega}{\pi \hbar \, 2^{2(n+1)} \, [(n+1)\,!]^2} \right)^{1/4} \left(y - \frac{d}{dy} \right)^{n+1} e^{-\frac{y^2}{2}} \tag{6.79}$$

$$= \left(\frac{m\omega}{\pi \hbar \, 2^{2(n+1)} \, [(n+1)\,!]^2} \right)^{1/4} \left(y - \frac{d}{dy} \right) H_n(y) \, e^{-\frac{y^2}{2}} \tag{6.80}$$

$$= \left(\frac{m\omega}{\pi \hbar \, 2^{2(n+1)} \, [(n+1)\,!]^2} \right)^{1/4} \left(2\, y \, H_n(y) - \frac{d}{dy} \, H_n(y) \right) e^{-\frac{y^2}{2}} \tag{6.81}$$

$$= \left(\frac{m\omega}{\pi \hbar \, 2^{2(n+1)} \, [(n+1)\,!]^2} \right)^{1/4} H_{n+1}(y) \, e^{-\frac{y^2}{2}} = \psi_{E_{n+1},x}(y) \tag{6.82}$$

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 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:

$$\langle x|X|x'\rangle = x\delta(x - x') \qquad \langle x|P|x'\rangle = -i\hbar\frac{d}{dx}\delta(x - x') \qquad (6.83)$$

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

$$\Omega(X, P) = \omega(x \to X, p \to P) \qquad (6.84)$$

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.

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*

$$[X, P] = i\,\hbar\, I \tag{6.85}$$

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

$$\Omega(X, P) = \omega(x \to X, p \to P) \tag{6.86}$$

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

The obvious question is: is the above version consistent with our original version of the postulate? The answer is yes, and we can see this as follows. Let us 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

$$X|x\rangle = x|x\rangle \tag{6.87}$$

If we normalize $\{|x\rangle\}$ such that $\langle x|x'\rangle = \delta(x - x')$, then we have

$$\langle x|X|x'\rangle = x\,\delta(x - x') \tag{6.88}$$

With that statement and the requirement $[X, P] = i\hbar I$, what must $\langle x | P | x' \rangle$ be? Let's just calculate it. As we proceed through it, we will see that the calculation is much like the one we did in Section 3.9 to calculate the matrix elements of the D operator in the $\{|x\rangle\}$ basis.

$$
\begin{aligned}
\langle x | [X, P] | x' \rangle &= \langle x | X P | x' \rangle - \langle x | P X | x' \rangle \\
&= \int_{-\infty}^{\infty} dx'' \left[\langle x | X | x'' \rangle \langle x'' | P | x' \rangle - \langle x | P | x'' \rangle \langle x'' | X | x' \rangle \right] \\
&= \int_{-\infty}^{\infty} dx'' \left[x\, \delta(x - x'')\, \langle x'' | P | x' \rangle - \langle x | P | x'' \rangle\, x'' \delta(x'' - x') \right]
\end{aligned}
$$
(6.89)

This is where things get tricky. We don't know what $\langle x'' | P | x' \rangle$ and $\langle x | P | x'' \rangle$ are, so we don't know that it is ok to just use the delta functions and integrate over x''. In fact, it is clear that if we just blindly do that, the right side will vanish and the commutator will not be what we want it to be. But how can we extract the matrix elements of P from the integrals to get explicit formulae for them?

We need to remember that our only aim here is to check consistency. So let's just do that, inserting $\langle x\,|P|x'\,\rangle = -i\,\hbar\,\frac{d}{dx}\,\delta(x-x')$ and seeing if we get the expected result:

$$
\begin{aligned}
\langle x\,|[X,P]\,|x'\,\rangle &= -i\,\hbar \int_{-\infty}^{\infty} dx''\Bigg\{ x\,\delta(x-x'')\left[\frac{d}{dx''}\,\delta(x''-x')\right] \\
&\qquad\qquad - \left[\frac{d}{dx}\,\delta(x-x'')\right] x''\,\delta(x''-x')\Bigg\} \\
&= -i\,\hbar \int_{-\infty}^{\infty} dx''\Bigg\{ x\,\delta(x-x'')\left[\frac{d}{dx''}\,\delta(x''-x')\right] \\
&\qquad\qquad - \left[\delta(x-x'')\frac{d}{dx''}\right] x''\,\delta(x''-x')\Bigg\} \\
&= -i\,\hbar \int_{-\infty}^{\infty} dx''\Bigg\{ x\,\delta(x-x'')\left[\frac{d}{dx''}\,\delta(x''-x')\right] \\
&\qquad - \delta(x-x'')\,\delta(x''-x') - \delta(x-x'')\,x''\left[\frac{d}{dx''}\,\delta(x''-x')\right]\Bigg\}
\end{aligned}
$$

$$(6.90)$$

where we have just used generic properties of the derivative of the delta function.

Now, $x\,\delta(x - x'') = x''\delta(x - x'')$ because, if one integrates either with a test function, one gets the same result:

$$\int_{-\infty}^{\infty} dx''\, x\, \delta(x - x'')\, f(x'') = x\, f(x) \tag{6.91}$$

$$\int_{-\infty}^{\infty} dx''\, x''\delta(x - x'')\, f(x'') = x\, f(x) \tag{6.92}$$

Therefore, the first term and the third term cancel, leaving

$$\langle x\, |[X, P]\, |x'\,\rangle = i\,\hbar \int_{-\infty}^{\infty} dx''\delta(x - x'')\,\delta(x'' - x') = i\,\hbar\,\delta(x - x') \tag{6.93}$$

as desired.

It is illustrative to discuss the many ways in which the above proof can go awry by mishandling the delta functions. Consider the following:

$$
\begin{aligned}
\langle x | [X, P] | x' \rangle &= -i\hbar \int_{-\infty}^{\infty} dx'' \Bigg\{ x\, \delta(x - x'') \left[\frac{d}{dx''} \delta(x'' - x') \right] \\
&\qquad\qquad - \left[\frac{d}{dx} \delta(x - x'') \right] x''\, \delta(x'' - x') \Bigg\} \\
&= i\hbar \left\{ x \left[\frac{d}{dx} \delta(x - x') \right] - \left[\frac{d}{dx} \delta(x - x') \right] x' \right\} \\
&= i\hbar \left\{ x \left[\frac{d}{dx} \delta(x - x') \right] - \left[\delta(x - x') \frac{d}{dx'} \right] x' \right\} \\
&= i\hbar \left\{ x \left[\frac{d}{dx} \delta(x - x') \right] - \delta(x - x') \right\}
\end{aligned}
\tag{6.94}
$$

That's not what we expect. What did we do wrong? In the first step, we assumed that we could treat derivatives of delta functions like other functions when multiplied against a delta function, and just replace the variables as one would expect. We have never proven that!

Rewriting Postulate 2 (cont.)

Another incorrect way to do the proof is:

$$
\begin{aligned}
\langle x \,|[X,P]\,|x'\rangle &= -i\,\hbar \int_{-\infty}^{\infty} dx'' \Bigg\{ x\,\delta(x-x'') \left[\frac{d}{dx''}\,\delta(x''-x') \right] \\
&\qquad\qquad\qquad - \left[\frac{d}{dx}\,\delta(x-x'') \right] x''\,\delta(x''-x') \Bigg\} \\
&= i\,\hbar \int_{-\infty}^{\infty} dx'' \Bigg\{ x\,\delta(x-x'') \left[\frac{d}{dx'}\,\delta(x''-x') \right] \\
&\qquad\qquad\qquad - \frac{d}{dx} \left[\delta(x'-x'')\,x''\,\delta(x''-x') \right] \Bigg\} \\
&= i\,\hbar \int_{-\infty}^{\infty} dx'' \Bigg\{ \frac{d}{dx'} \left[x\,\delta(x-x'')\,\delta(x''-x') \right] \\
&\qquad\qquad\qquad - \frac{d}{dx} \left[\delta(x'-x'')\,x''\,\delta(x''-x') \right] \Bigg\}
\end{aligned}
\tag{6.95}
$$

$$\langle x \,|[X, P]\,|x\,'\,\rangle = i\,\hbar \,\frac{d}{dx\,'} \int_{-\infty}^{\infty} dx\,'' \left[x\,\delta(x - x\,'')\,\delta(x\,'' - x\,') \right]$$

$$- i\,\hbar \,\frac{d}{dx} \int_{-\infty}^{\infty} dx\,'' \left[\delta(x\,' - x\,'')\,x\,''\,\delta(x\,'' - x\,') \right]$$

$$= i\,\hbar \,\frac{d}{dx\,'} \left[x\,\delta(x - x\,') \right] - i\,\hbar \,\frac{d}{dx} \left[x\,'\delta(x - x\,') \right] \qquad (6.96)$$

$$= i\,\hbar\,x \left[\frac{d}{dx\,'}\,\delta(x - x\,') \right] - i\,\hbar\,x\,' \left[\frac{d}{dx}\,\delta(x - x\,') \right]$$

$$= -i\,\hbar\,x \left[\frac{d}{dx}\,\delta(x - x\,') \right] - i\,\hbar\,x\,' \left[\frac{d}{dx}\,\delta(x - x\,') \right]$$

which again is not what we want. Here, the error was to assume that we could bring the derivative outside the integral when the derivative is being taken with respect to a variable that none of the other factors depend on. This would be fine for normal functions, but, because delta functions tie all the variables together, that step was not valid. Another way of looking at it is that we never proved that one could move a derivative that acts on a delta function outside the integral.

There are probably other failure modes. One has to be very careful to never assume a property about the delta function that one has not explicitly proven.

Shankar points out that we could have added an arbitrary function of x to our definition of P:

$$\langle x \,|P|x'\rangle = -i\,\hbar\left[\frac{d}{dx}\,\delta(x-x')\right] + f(x)\,\delta(x-x') \qquad (6.97)$$

Since $f(x)$ commutes with x and x', adding this term would not change the commutator. However, we stick with the simplest definition, with $f(x) = 0$.

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 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

$$[X_i, P_j] = i\,\hbar\,\delta_{ij} \qquad [X_i, X_j] = 0 \qquad [P_i, P_j] = 0 \qquad (6.98)$$

Any arbitrary classical dynamical variable $\omega\left(\{x_i\}, \{p_i\}\right)$ has a corresponding Hermitian operator

$$\Omega\left(\{X_i\}, \{P_i\}\right) = \omega\left(\{x_i \to X_i\}, \{p_i \to P_i\}\right) \qquad (6.99)$$

where we simply replace $\{x_i\}$ and $\{p_i\}$ in ω with $\{X_i\}$ and $\{P_i\}$ to obtain $\Omega\left(\{X_i\}, \{P_i\}\right)$.

To be utterly clear, the original writing of the *N*-dimensional version of Postulate 2 would tell us that the first commutator is

$$
\begin{aligned}
\langle x_1, x_2, \ldots, x_N \,|[X_i, P_j]|x_1', x_2', \ldots, x_N' \rangle \\
= i\,\hbar\,\delta_{ij}\,\delta(x_1 - x_1')\,\delta(x_2 - x_2') \cdots \delta(x_N - x_N')
\end{aligned} \quad (6.100)
$$

This is consistent with our new writing as follows:

$$
\begin{aligned}
\langle x_1, x_2, \ldots, x_N \,|[X_i, P_j]|x_1', x_2', \ldots, x_N' \rangle &= \langle x_1, x_2, \ldots, x_N \,|i\,\hbar\,\delta_{ij}|x_1', x_2', \ldots, x_N' \rangle \\
&= i\,\hbar\,\delta_{ij}\,\delta(x_1 - x_1')\,\delta(x_2 - x_2') \cdots \delta(x_N - x_N')
\end{aligned}
$$
$$(6.101)$$

(The other commutators always vanish, so we do not bother to write them out to the above level of details.) The way we build up the Hilbert space for *N* particles from the single-particle Hilbert spaces is somewhat subtle and we will go through it carefully later.

We will generalize Postulate 2 again in a similar manner when we preview quantum field theory.

Section 7
The Heisenberg Uncertainty Relation

Lecture 19:
The Heisenberg Uncertainty Relation

Revision Date: 2007/11/12

Deriving the Uncertainty Relation

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

$$\langle \Omega \rangle = \langle \psi \,|\Omega|\psi \,\rangle = \sum_i P(\omega_i)\, \omega_i \quad \text{or} \quad \int_{\omega_-}^{\omega_+} d\omega \, P(\omega)\, \omega \tag{7.1}$$

$$\langle (\Delta\Omega)^2 \rangle = \langle \psi \,|\, [\Omega - \langle \Omega \rangle]^2 \,|\psi \,\rangle = \langle \psi \,|\, \left[\Omega^2 - \langle \Omega \rangle^2\right] |\psi \,\rangle = \langle \psi \,|\Omega^2|\psi \,\rangle - \langle \Omega \rangle^2 \tag{7.2}$$

$$= \sum_i P(\omega_i)\,(\omega_i - \langle \Omega \rangle)^2 \quad \text{or} \quad \int_{\omega_-}^{\omega_+} d\omega \, P(\omega)\,(\omega - \langle \Omega \rangle)^2 \tag{7.3}$$

Deriving the Uncertainty Relation (cont.)

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:

$$[\Omega, \Lambda] = i\,\Gamma \tag{7.4}$$

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).

Now, consider the product of the squares of the uncertainties of the two operators for an arbitrary state $|\psi\rangle$:

$$
\begin{aligned}
\langle(\Delta\Omega)^2\rangle\langle(\Delta\Lambda)^2\rangle &= \langle\psi\,|\widetilde{\Omega}^2|\psi\rangle\langle\psi\,|\widetilde{\Lambda}^2|\psi\rangle \quad \text{with} \quad \widetilde{\Omega} = \Omega - \langle\Omega\rangle,\ \widetilde{\Lambda} = \Lambda - \langle\Lambda\rangle \\
&= \left(\widetilde{\Omega}^\dagger|\psi\rangle\right)^\dagger \widetilde{\Omega}|\psi\rangle \left(\widetilde{\Lambda}^\dagger|\psi\rangle\right)^\dagger \widetilde{\Lambda}|\psi\rangle \\
&= \left(\widetilde{\Omega}|\psi\rangle\right)^\dagger \widetilde{\Omega}|\psi\rangle \left(\widetilde{\Lambda}|\psi\rangle\right)^\dagger \widetilde{\Lambda}|\psi\rangle \quad \text{bec.}\ \widetilde{\Omega},\ \widetilde{\Lambda}\ \text{are Hermitian} \\
&= \left|\widetilde{\Omega}|\psi\rangle\right|^2 \left|\widetilde{\Lambda}|\psi\rangle\right|^2 \\
&\geq \left|\left(\widetilde{\Omega}|\psi\rangle\right)^\dagger \widetilde{\Lambda}|\psi\rangle\right|^2 \quad \text{by Schwarz Inequality, Equation 3.15} \qquad (7.5) \\
&= \left|\langle\psi\,|\widetilde{\Omega}\,\widetilde{\Lambda}\,|\psi\rangle\right|^2 \\
&= \left|\langle\psi\,|\left\{\frac{1}{2}\left[\widetilde{\Omega}\,\widetilde{\Lambda} + \widetilde{\Lambda}\,\widetilde{\Omega}\right] + \frac{1}{2}\left[\widetilde{\Omega}\,\widetilde{\Lambda} - \widetilde{\Lambda}\,\widetilde{\Omega}\right]\right\}|\psi\rangle\right|^2 \\
&= \left[\frac{1}{2}\langle\psi\,|\left[\widetilde{\Omega},\widetilde{\Lambda}\right]_+ |\psi\rangle + \frac{1}{2}\langle\psi\,|\left[\widetilde{\Omega},\widetilde{\Lambda}\right]|\psi\rangle\right]^2
\end{aligned}
$$

To evaluate the above, we need to recognize 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

$$\langle(\Delta\Omega)^2\rangle\langle(\Delta\Lambda)^2\rangle \geq \frac{1}{4}\left[\langle\psi\,|\,\left[\widetilde{\Omega}, \widetilde{\Lambda}\right]_+ |\psi\,\rangle\right]^2 + \frac{1}{4}\left[\langle\psi\,|\Gamma|\psi\,\rangle\right]^2 \tag{7.6}$$

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

$$\langle(\Delta\Omega)^2\rangle\langle(\Delta\Lambda)^2\rangle \geq \frac{1}{4}\left[\langle\psi\,|\,\left[\widetilde{\Omega}, \widetilde{\Lambda}\right]_+ |\psi\,\rangle\right]^2 + \frac{\hbar^2}{4} \tag{7.7}$$

$$\text{or} \qquad \sqrt{\langle(\Delta\Omega)^2\rangle}\sqrt{\langle(\Delta\Lambda)^2\rangle} \geq \frac{\hbar}{2} \tag{7.8}$$

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

$$\widetilde{\Omega}|\psi\rangle = c\,\widetilde{\Lambda}\,|\psi\rangle \tag{7.9}$$

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:

$$\langle\psi\,|\left[\widetilde{\Omega},\widetilde{\Lambda}\right]_{+}|\psi\rangle = 0 \tag{7.10}$$

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

$$\sqrt{\langle(\Delta X)^2\rangle}\sqrt{\langle(\Delta P)^2\rangle} = \frac{\hbar}{2} \tag{7.11}$$

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,

$$(P - \langle P\rangle)\,|\psi\rangle = c\,(X - \langle X\rangle)\,|\psi\rangle \tag{7.12}$$

$$\left(-i\,\hbar\,\frac{d}{dx} - \langle P\rangle\right)\psi(x) = c\,(x - \langle X\rangle)\,\psi(x) \tag{7.13}$$

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.

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.

Examples (cont.)

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

$$\sqrt{\langle (\Delta P_y)^2 \rangle} \geq \frac{\hbar}{2} \frac{1}{\sqrt{\langle (\Delta Y)^2 \rangle}} = \frac{\hbar \sqrt{3}}{2a} \tag{7.14}$$

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

$$\sqrt{\langle (\Delta \theta)^2 \rangle} = \frac{\sqrt{\langle (\Delta P_y)^2 \rangle}}{\langle P_x \rangle} \geq \frac{\frac{\hbar \sqrt{3}}{2a}}{\hbar k} = \frac{\sqrt{3}}{2ka} \tag{7.15}$$

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

$$\frac{\langle (\Delta P)^2 \rangle}{2\,m_p} = \alpha \quad \Longrightarrow \quad \langle (\Delta P)^2 \rangle = 2\,\alpha\,m_p \tag{7.16}$$

From this, let's use the uncertainty principle to determine the position uncertainty

$$\langle (\Delta X)^2 \rangle = \frac{\hbar^2}{4} \frac{1}{\langle (\Delta P)^2 \rangle} = \frac{\hbar^2}{4} \frac{1}{2\,\alpha\,m_p} \tag{7.17}$$

Numerically, we have

$$\sqrt{\langle (\Delta X)^2 \rangle} = \frac{\hbar}{2\,A\,\sqrt{2\,\alpha\,m_p}} \tag{7.18}$$

$$= \frac{1.0 \times 10^{-34}\ \text{J s}}{2\,A\,\sqrt{2\,\alpha \times 10^6 \times 1.6 \times 10^{-19}\ \text{J} \times 1.7 \times 10^{-27}\ \text{kg}}} \tag{7.19}$$

$$= \frac{2.1 \times 10^{-15}\ \text{m}}{\sqrt{\alpha}} \tag{7.20}$$

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 20:
The Energy-Time Uncertainty Relation
Multiparticle Systems: Setup

Revision Date: 2007/11/14

The Energy-Time Uncertainty Relation

H and $i\hbar \frac{d}{dt}$ as conjugate operators

The Schrödinger Equation tells us

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle \tag{7.21}$$

Let's take the matrix element on the left with $\langle x |$ and insert completeness in the $\{|x\rangle\}$-basis, and integrate over the completeness relation to obtain

$$i\hbar \frac{d}{dt} \psi_x(x, t) = \int_{-\infty}^{\infty} dx'\langle x |H| x'\rangle \psi_x(x', t) \tag{7.22}$$

for the one-dimensional problems we have considered. Notice the striking similarity to the expression for the P operator acting in the $\{|x\rangle\}$ basis:

$$-i\hbar \frac{d}{dx} \psi_x(x, t) = \int_{-\infty}^{\infty} dx'\langle x |P| x'\rangle \psi_x(x', t) \tag{7.23}$$

The variables t and E are mathematically conjugate to each other in a manner similar to x and p. However, it must be emphasized that this is not rigorous in the sense of operators because there is no t operator!

We can see the conjugate relationship more rigorously by considering a free-particle wave packet with

$$\langle X \rangle = \frac{p_0}{m}\, t \qquad \langle (\Delta X)^2 \rangle_{t=0} = \sigma_x^2 \qquad \langle P \rangle = p_0 \qquad \langle (\Delta P)^2 \rangle = \sigma_p^2 = \frac{\hbar^2}{4\,\sigma_x^2} \qquad (7.24)$$

with $\sigma_p/p \ll 1$. The wavefunction as a function of position and time is

$$\psi_x(x,t) = \langle x\,|\psi(t)\,\rangle = \langle x\,|U(t)\,|\psi(t=0)\,\rangle \qquad (7.25)$$

$$= \int_0^\infty dp\, e^{-\frac{i}{\hbar}\frac{p^2}{2m}t}\, [\langle x\,|E_p\,\rangle\langle E_p\,|\psi(t=0)\,\rangle + \langle x\,|E_{-p}\,\rangle\langle E_{-p}\,|\psi(t=0)\,\rangle] \qquad (7.26)$$

$$\approx \int_0^\infty dp\, e^{-\frac{i}{\hbar}\frac{p^2}{2m}t}\, e^{\frac{i}{\hbar}px}\psi_p(p) \qquad (7.27)$$

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 $p < 0$.

Now, let's specialize to $x = 0$: we want to see what the time dependence of the wavefunction at the origin is. We have

$$\psi_x(0, t) \approx \int_0^\infty dp \, e^{-\frac{i}{\hbar} \frac{p^2}{2m} t} \psi_p(p) \tag{7.28}$$

$$\approx \int_{-\infty}^\infty dE \, e^{-\frac{i}{\hbar} E t} \psi_E(E) \tag{7.29}$$

$$\text{where} \quad \psi_E(E) = \sqrt{\frac{m}{2E}} \, \psi_p(p) = \frac{m}{p} \, \psi_p(p) \tag{7.30}$$

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.

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

$$\sqrt{\langle (\Delta E)^2 \rangle} \approx 2\,E\,\frac{\sigma_p}{p_0} = \frac{p_0^2}{m}\,\frac{\sigma_p}{p_0} = \frac{p_0}{m}\,\sigma_p \qquad (7.31)$$

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:

$$\sqrt{(\Delta t)^2} = \frac{\sqrt{\langle (\Delta X)^2 \rangle}}{v} = \frac{\sigma_x(t)}{p_0/m} \geq \frac{\sigma_x}{p_0/m} \qquad (7.32)$$

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.

Multiplying the two uncertainties together, we have

$$\sqrt{\langle(\Delta E)^2\rangle}\sqrt{(\Delta t)^2} \geq \frac{p_0}{m}\,\sigma_p\frac{\sigma_x}{p_0/m} = \sigma_p\,\sigma_x = \frac{\hbar}{2} \tag{7.33}$$

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.

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.

▶ **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
Multiparticle Systems

Direct Product Spaces

Direct Product Spaces

Suppose we have two Hilbert spaces, $\mathcal{V}_1$ and $\mathcal{V}_2$, each containing the states corresponding to a particular degree of freedom; a typical example is that $\mathcal{V}_1$ contains the states for particle 1 and $\mathcal{V}_2$ for particle 2, where both particles live in a single spatial dimension. Then we can build a new Hilbert space, $\mathcal{V} = \mathcal{V}_1 \otimes \mathcal{V}_2$, 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:

$$|v,\, w\,\rangle^{(1)\otimes(2)} \equiv |v\,\rangle^{(1)} \otimes |w\,\rangle^{(2)} \tag{8.1}$$

where the superscripts on each ket outside the ket bracket indicates which particle's Hilbert space it belongs to: the $^{(1)}$ kets belong to $\mathcal{V}_1$, the $^{(2)}$ kets to $\mathcal{V}_2$ and the $^{(1)\otimes(2)}$ kets to $\mathcal{V} = \mathcal{V}_1 \otimes \mathcal{V}_2$. 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, $|\hat{x}\,\rangle$, and for $|v\,\rangle^{(2)}$ to be a basis element of the position-basis element for particle 2, $|\tilde{x}\,\rangle$. The direct product vector $|\hat{x}, \tilde{x}\,\rangle^{(1)\otimes(2)} = |\hat{x}\,\rangle^{(1)} \otimes |\tilde{x}\,\rangle^{(2)}$ 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 $|x_1\rangle^{(1)}$, $|x_2\rangle^{(2)}$ and $|x_1, x_2\rangle^{(1)\otimes(2)}$, and possibly dropped the superscripts altogether. There is nothing wrong with this, but the use of the numbered subscripts somehow 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. But the particles live in the same physical space, both x_1 and x_2 are accessible to both, it would be perfectly reasonable to have $|x_2\rangle^{(1)}$, $|x_1\rangle^{(2)}$, and $|x_2, x_1\rangle^{(1)\otimes(2)} = |x_2\rangle^{(1)} \otimes |x_1\rangle^{(2)}$. That's confusing, but it is convention and we shall soon collapse to using that kind of notation.

▶ Second, we allow linear combinations of elements of the above type:

$$|u\rangle^{(1)\otimes(2)} = \left(\alpha_1|v\rangle^{(1)} + \alpha_2|w\rangle^{(1)}\right) \otimes \beta|v\rangle^{(2)} \tag{8.2}$$

$$= \alpha_1\beta\left(|v\rangle^{(1)} \otimes |v\rangle^{(2)}\right) + \alpha_2\beta\left(|w\rangle^{(1)} \otimes |v\rangle^{(2)}\right) \tag{8.3}$$

α_1, α_2, and β must belong to the same scalar field in order for there to be a single scalar field for the direct product space.

▶ Third, the inner product is defined as the obvious extension of the inner product in each space:

$$^{(1)\otimes(2)}\langle v_1, w_1 \,|\, v_2, w_2 \rangle^{(1)\otimes(2)} \equiv \left({}^{(1)}\langle v_1 \,|\, v_2 \rangle^{(1)} \right) \left({}^{(2)}\langle w_1 \,|\, w_2 \rangle^{(2)} \right) \qquad (8.4)$$

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 $\{|n\rangle^{(2)}\}$ are a basis for the second space. Then a basis for the direct product space consists of all products of the form

$$|n, m\rangle^{(1)\otimes(2)} = |n\rangle^{(1)} \otimes |m\rangle^{(2)} \qquad (8.5)$$

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 $N \times M$.

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.

Direct Product Spaces (cont.)

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:

$$|0\rangle^{(1)} \otimes |w\rangle^{(2)} = |0\rangle^{(1)\otimes(2)} = |v\rangle^{(1)} \otimes |0\rangle^{(2)} \tag{8.6}$$

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*.

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:

$$-|v, w\rangle^{(1)\otimes(2)} = -\left(|v\rangle^{(1)} \otimes |w\rangle^{(2)}\right) \tag{8.7}$$

$$= \left(-|v\rangle^{(1)}\right) \otimes |w\rangle^{(2)} = |v\rangle^{(1)} \otimes \left(-|w\rangle^{(2)}\right) \tag{8.8}$$

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:

$$|n_1\rangle^{(1)} \otimes |m_1\rangle^{(2)} + |n_2\rangle^{(1)} \otimes |m_2\rangle^{(2)} \neq |n\rangle^{(1)} \otimes |m\rangle^{(2)} \quad \text{for all } |n\rangle^{(1)}, |m\rangle^{(2)} \tag{8.9}$$

It is easy to prove this by assuming the left side can be written as a simple direct product, expanding both sides in terms of their respective spaces basis elements, and seeing that it is impossible for the coefficients to be equal. 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 Product Spaces (cont.)

Pitfalls in Understanding Direct Product Spaces

Perhaps the biggest problem in understanding direct product states is the desire to want to put the particles in the same Hilbert space because they live in the same spatial 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 probability of detecting the two particles 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-spatial-dimensional single-particle 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-particle 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 a 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.

Multiparticle Systems: Direct Product Spaces

Direct Product Spaces (cont.)

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 $\{|n_i\rangle^{(1)}\}$ as the basis for $\mathcal{V}_1$ and $\{|m_i\rangle^{(2)}\}$ as the basis for $\mathcal{V}_2$. Let $|v\rangle^{(1)}$ and $|w\rangle^{(2)}$ be vectors in the two spaces. Then we have

$$|v\rangle^{(1)} = \sum_i v_i |n_i\rangle^{(1)} \qquad |w\rangle^{(2)} = \sum_j w_j |m_j\rangle^{(2)} \tag{8.10}$$

$$|v\rangle^{(1)} \otimes |w\rangle^{(2)} = \left(\sum_i v_i |n_i\rangle^{(1)}\right) \otimes \left(\sum_j w_j |m_j\rangle^{(2)}\right) \tag{8.11}$$

$$= \sum_{i,j} v_i\, w_j \left(|n_i\rangle^{(1)} \otimes |m_j\rangle^{(2)}\right) = \sum_{i,j} v_i\, w_j \left(|n_i, w_j\rangle^{(1)\otimes(2)}\right) \tag{8.12}$$

Direct Product Spaces (cont.)

However, the above state is not the most generic possible state in the direct product space; that would be

$$|u\rangle^{(1)\otimes(2)} = \sum_{i,j} u_{ij} |v_i, w_j\rangle^{(1)\otimes(2)} \tag{8.13}$$

For arbitrary $\{u_{ij}\}$, one cannot decompose u_{ij} in the form $u_{ij} = v_i w_j$. In particular, let's consider the example from earlier:

$$|u\rangle^{(1)\otimes(2)} = |n_1\rangle^{(1)} \otimes |m_1\rangle^{(2)} + |n_2\rangle^{(1)} \otimes |m_2\rangle^{(2)} \tag{8.14}$$

We have

$$u_{11} = 0 \qquad u_{12} = 1 \qquad u_{21} = 1 \qquad u_{22} = 0 \tag{8.15}$$

Let's assume an expansion $u_{ij} = v_i w_j$. The statement $u_{11} = 0$ implies that either $v_1 = 0$ or $w_1 = 0$. But then one of $u_{12} = v_1 w_2$ or $u_{21} = v_2 w_1$ must vanish; they do not. Contradiction.

Direct Product Spaces (cont.)

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 particle 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 $\mathcal{V} = \mathcal{V}_1 \otimes \mathcal{V}_2$, an operator $\Omega^{(1)}$ that acts in $\mathcal{V}_1$ simply does not exist in $\mathcal{V}_2$. The natural thing to do, then, is to assign it the action of the identity operator in $\mathcal{V}_2$, $I^{(2)}$. 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

$$\Omega^{(1)\otimes(2)} = \Omega^{(1)} \otimes I^{(2)} \tag{8.16}$$

or, explicitly,

$$\Omega^{(1)\otimes(2)} |v, w\rangle^{(1)\otimes(2)} = \left(\Omega^{(1)} |v\rangle^{(1)}\right) \otimes \left(I^{(2)} |w\rangle^{(2)}\right) \tag{8.17}$$

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

$$X_1 |\widehat{x}\,\rangle^{(1)} = \widehat{x}\,|\widehat{x}\,\rangle^{(1)} \qquad X_2 |\widetilde{x}\,\rangle^{(2)} = \widetilde{x}\,|\widetilde{x}\,\rangle^{(2)} \tag{8.18}$$

$$X_1 \left(|\widehat{x}\,\rangle^{(1)} \otimes |\widetilde{x}\,\rangle^{(2)} \right) = \left(X_1 |\widehat{x}\,\rangle^{(1)} \right) \otimes \left(I |\widetilde{x}\,\rangle^{(2)} \right) = \widehat{x} \left(|\widehat{x}\,\rangle^{(1)} \otimes |\widetilde{x}\,\rangle^{(2)} \right) \tag{8.19}$$

$$X_2 \left(|\widehat{x}\,\rangle^{(1)} \otimes |\widetilde{x}\,\rangle^{(2)} \right) = \left(I |\widehat{x}\,\rangle^{(1)} \right) \otimes \left(X_2 |\widetilde{x}\,\rangle^{(2)} \right) = \widetilde{x} \left(|\widehat{x}\,\rangle^{(1)} \otimes |\widetilde{x}\,\rangle^{(2)} \right) \tag{8.20}$$

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_i and P_i for a direct product space consisting of N factors:

$$\left\langle \prod_{k=1}^{N} x_k \left| X_i^{\prod_{n=1}^{N}(n)} \right| \prod_{q=1}^{N} x_q' \right\rangle = \langle x_i | X_i^{(i)} | x_i' \rangle \prod_{n \neq i}^{N} \langle x_n | I^{(n)} | x_n' \rangle \tag{8.21}$$

$$= x_i \, \delta(x_i - x_i') \prod_{n \neq i}^{N} \delta(x_n - x_n') = x_i \prod_{n=1}^{N} \delta(x_n - x_n') \tag{8.22}$$

$$\left\langle \prod_{k=1}^{N} x_k \left| P_i^{\prod_{n=1}^{N}(n)} \right| \prod_{q=1}^{N} x_q' \right\rangle = \langle x_i | P_i^{(i)} | x_i' \rangle \prod_{n \neq i}^{N} \langle x_n | I^{(n)} | x_n' \rangle \tag{8.23}$$

$$= \left(-i \hbar \frac{d}{dx_i} \delta(x_i - x_i') \right) \prod_{n \neq i}^{N} \delta(x_n - x_n') \tag{8.24}$$

Or, if we want to check the commutators we postulated in Section 6.5, we have

$$\left\langle \prod_{k=1}^{N} x_k \left| \left[X_i^{\prod_{q=1}^{N}(q)}, P_i^{\prod_{m=1}^{N}(m)} \right] \right| \prod_{n=1}^{N} x_n' \right\rangle$$

$$= \langle x_i | \left[X_i^{(i)}, P_i^{(i)} \right] | x_i' \rangle \prod_{n \neq i}^{N} \langle x_n | I^{(n)} | x_n' \rangle = i\,\hbar\,\delta(x_i - x_i') \prod_{n \neq i}^{N} \delta(x_n - x_n')$$

$$= i\,\hbar \prod_{n=1}^{N} \delta(x_n - x_n') \tag{8.25}$$

and, for $i \neq j$,

$$
\left\langle \prod_{k=1}^{N} x_k \left| \left[X_i^{\prod_{q=1}^{N}(q)}, P_j^{\prod_{m=1}^{N}(m)} \right] \right| \prod_{n=1}^{N} x_n' \right\rangle
$$

$$
= \langle x_i | \left[X_i^{(i)}, I^{(i)} \right] | x_i' \rangle \langle x_j | \left[I^{(j)}, P_j^{(j)} \right] | x_j' \rangle \prod_{n \neq i,j}^{N} \langle x_n | I^{(n)} | x_n' \rangle = (0)(0) \prod_{n \neq i,j}^{N} \langle x_n | I^{(n)} | x_n' \rangle
$$

$$
= 0 \tag{8.26}
$$

One could prove $\left[X_i^{\prod_{q=1}^{N}(q)}, X_j^{\prod_{m=1}^{N}(m)} \right] = 0$ and $\left[P_i^{\prod_{q=1}^{N}(q)}, P_j^{\prod_{m=1}^{N}(m)} \right] = 0$ by a similar technique.

Lecture 21:
Multiparticle Systems: The Hamiltonian and Time Evolution
Position-Space Wavefunction

Revision Date: 2007/11/16

The Hamiltonian and Time-Evolution

Time Evolution

For a direct product space $V = \prod_{n=1}^{N} V_n$, there is a direct product space Hamiltonian operator $H^{\prod_{n=1}^{N}}$ that gives the time evolution via the Schrödinger Equation:

$$i\,\hbar\,\frac{d}{dt}\,|\psi\,\rangle^{\prod_{n=1}^{N}(n)} = H^{\prod_{m=1}^{N}(m)}\,|\psi\,\rangle^{\prod_{n=1}^{N}(n)} \tag{8.27}$$

As usual, $H^{\prod_{m=1}^{N}(m)}$ will have eigenstates that satisfy

$$H^{\prod_{m=1}^{N}(m)}|\psi_E\,\rangle^{\prod_{n=1}^{N}(n)} = E\,|\psi_E\,\rangle^{\prod_{n=1}^{N}(n)} \tag{8.28}$$

For these energy eigenstates, the Schrödinger Equation is

$$i\,\hbar\,\frac{d}{dt}\,|\psi_E\,\rangle^{\prod_{n=1}^{N}(n)} = E\,|\psi_E\,\rangle^{\prod_{n=1}^{N}(n)} \tag{8.29}$$

and the time evolving solution is

$$|\psi_E(t)\,\rangle^{\prod_{n=1}^{N}(n)} = e^{-\frac{i}{\hbar}\,E\,t}|\psi_E\,\rangle^{\prod_{n=1}^{N}(n)} \tag{8.30}$$

Separable Hamiltonians

A separable Hamiltonian is one that can be written in the form

$$H^{\Pi_{n=1}^{N}(n)} = \sum_{n=1}^{N} H^{(n)} \otimes \prod_{m \neq n} I^{(m)} \tag{8.31}$$

The eigenvalue-eigenvector equation simplifies to a set of single-particle Schrödinger equations in this case. We can see this as follows. First, write down the eigenvalue-eigenvector equation:

$$E \, |\psi_E\rangle^{\Pi_{q=1}^{N}(q)} = H^{\Pi_{n=1}^{N}(n)} |\psi_E\rangle^{\Pi_{q=1}^{N}(q)} = \sum_{n=1}^{N} \left(H^{(n)} \otimes \prod_{m \neq n} I^{(m)} \right) |\psi_E\rangle^{\Pi_{q=1}^{N}(q)} \tag{8.32}$$

Now, expand the direct product space eigenstate in a basis consisting of direct products of eigenstates of the individual single particle Hamiltonians. We are assured that the latter are valid bases for the individual spaces 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 V_n that satisfy

$$H^{(n)} |\psi_{E_m^{(n)}}\rangle^{(n)} = E_m^{(n)} |\psi_{E_m^{(n)}}\rangle^{(n)} \tag{8.33}$$

where m tells us which eigenstate of $H^{(n)}$ we are referring to. Then our basis for the direct product space consists of direct product states of the form

$$\prod_{n=1}^{N} |\psi_{E_{m_n}^{(n)}}\rangle^{(n)} \tag{8.34}$$

The Hamiltonian and Time-Evolution (cont.)

Consider the action of the Hamiltonian on a state of this type:

$$H^{\prod_{n=1}^{N}(n)} |\psi_E\rangle^{\prod_{q=1}^{N}(q)} = \sum_{n=1}^{N} \left(H^{(n)} \otimes \prod_{q \neq n} I^{(q)} \right) \prod_{k=1}^{N} |\psi_{E_{m_k}^{(k)}}\rangle^{(k)} \tag{8.35}$$

$$= \sum_{n=1}^{N} \left[\left(H^{(n)} |\psi_{E_{m_n}^{(n)}}\rangle^{(n)} \right) \otimes \prod_{q \neq n} \left(I^{(q)} |\psi_{E_{m_q}^{(q)}}\rangle^{(q)} \right) \right] \tag{8.36}$$

$$= \sum_{n=1}^{N} E_{m_n}^{(n)} \left(\prod_{q=1}^{N} |\psi_{E_{m_q}^{(q)}}\rangle^{(q)} \right) \tag{8.37}$$

$$= \left(\sum_{n=1}^{N} E_{m_n}^{(n)} \right) \left(\prod_{q=1}^{N} |\psi_{E_{m_q}^{(q)}}\rangle^{(q)} \right) \tag{8.38}$$

$$\equiv E_{m_1 \cdots m_N} \left(\prod_{n=1}^{N} |\psi_{E_{m_n}^{(q)}}\rangle^{(n)} \right) \tag{8.39}$$

$$\text{with} \quad E_{m_1 \cdots m_N} = E_{m_1}^{(1)} + E_{m_2}^{(2)} + \cdots + E_{m_N}^{(N)} \tag{8.40}$$

The Hamiltonian and Time-Evolution (cont.)

The time evolution of the eigenstates of the direct product Hamiltonian are therefore

$$|\psi_{E_{m_1}\cdots m_N}(t)\rangle^{\prod_{n=1}^N (n)} = e^{-\frac{i}{\hbar} E_{m_1}\cdots m_N t}|\psi_{E_{m_1}\cdots m_N}\rangle^{\prod_{n=1}^N (n)} \tag{8.41}$$

$$= e^{-\frac{i}{\hbar} E_{m_1}\cdots m_N t}\left(\prod_{n=1}^N |\psi_{E_{m_n}^{(n)}}\rangle^{(n)}\right) \tag{8.42}$$

$$= \left(\prod_{n=1}^N e^{-\frac{i}{\hbar} E_{m_n}^{(n)} t}|\psi_{E_{m_n}^{(n)}}\rangle^{(n)}\right) \tag{8.43}$$

One just has the direct product of the time-evolving individual single-particle 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-particle 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 particle 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).

The Hamiltonian and Time-Evolution (cont.)

Non-Separable Hamiltonians

There are, of course, non-separable Hamiltonians. Usually, the non-separability comes from the potential energy term. A classic example is the Coulomb interaction between two particles. The Hamiltonian in the direct product space is

$$H = \frac{P_{1,x}^2 + P_{1,y}^2 + P_{1,z}^2}{2\,m_1} + \frac{P_{2,x}^2 + P_{2,y}^2 + P_{2,z}^2}{2\,m_2} \tag{8.44}$$
$$- \frac{e^2}{\sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2 + (Z_1 - Z_2)^2}}$$

where every operator is implied to have a $^{(1)\otimes(2)}$ superscript. 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

$$X_{CM} = \frac{m_1 X_1 + m_2 X_2}{M} \qquad P_{CM,x} = P_{1,x} + P_{2,x} \qquad (8.45)$$

$$X_{12} = X_1 - X_2 \qquad P_{12,x} = \frac{m_2}{M} P_{1,x} + \frac{m_1}{M} P_{2,x} \qquad M = m_1 + m_2 \qquad (8.46)$$

and so on for the y and z degrees of freedom. Then the Hamiltonian becomes

$$H = \frac{P_{CM,x}^2 + P_{CM,y}^2 + P_{CM,z}^2}{2\,M} + \frac{P_{12,x}^2 + P_{12,y}^2 + P_{12,z}^2}{2\,\mu} - \frac{e^2}{\sqrt{X_{12}^2 + Y_{12}^2 + Z_{12}^2}} \qquad (8.47)$$

$$\equiv H_{CM,x} + H_{CM,y} + H_{CM,z} + H_{12} \qquad (8.48)$$

$$\mu = \frac{m_1\,m_2}{M} \qquad (8.49)$$

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 somewhat separable. We will see later that H_{12} can be separated further by going to
spherical 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 Wavefunction

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:

$$|x_1, x_2, \cdots, x_N\rangle^{\prod_{n=1}^{N}(n)} = \prod_{n=1}^{N} |x_n\rangle^{(n)} \tag{8.50}$$

This is a complete basis for the direct product space because the factors are complete bases for the factor spaces.

Now, let's write the state $|\psi\rangle^{\Pi_{n=1}^{N}(n)}$ in terms of the position basis. That is, there are some set of coefficients for expanding $|\psi\rangle^{\Pi_{n=1}^{N}(n)}$ in terms of the $\{|x_1, x_2, \cdots, x_N\rangle^{\Pi_{n=1}^{N}(n)}\}$; there have to be, since the latter are a complete basis. Let's label those coefficients $\psi(x_1, x_2, \cdots, x_N)$:

$$
\begin{aligned}
&|\psi\rangle^{\Pi_{n=1}^{N}(n)} \\
&= \int_{-\infty}^{\infty} dx_1' \int_{-\infty}^{\infty} dx_2' \cdots \int_{-\infty}^{\infty} dx_N'\, \psi(x_1', x_2', \cdots, x_N')\, |x_1', x_2', \cdots, x_N'\rangle^{\Pi_{n=1}^{N}(n)} \\
&= \int_{-\infty}^{\infty} dx_1' \int_{-\infty}^{\infty} dx_2' \cdots \int_{-\infty}^{\infty} dx_N'\, \psi(x_1', x_2', \cdots, x_N') \prod_{n=1}^{N} |x_n'\rangle^{(n)}
\end{aligned}
\tag{8.51}
$$

Then, the position space wavefunction is obtained trivially:

$$
\begin{aligned}
&\Pi_{n=1}^{N}{}^{(n)}\langle x_1, x_2, \cdots, x_N \,|\psi\,\rangle \Pi_{n=1}^{N}{}^{(n)} \\
&= \left(\prod_{k=1}^{N} \langle x_k \,|^{(k)} \right) \int_{-\infty}^{\infty} dx_1' \int_{-\infty}^{\infty} dx_2' \cdots \int_{-\infty}^{\infty} dx_N' \qquad \psi(x_1', x_2', \cdots, x_N') \prod_{n=1}^{N} |x_n'\,\rangle^{(n)} \\
&= \int_{-\infty}^{\infty} dx_1' \int_{-\infty}^{\infty} dx_2' \cdots \int_{-\infty}^{\infty} dx_N' \qquad\quad \psi(x_1', x_2', \cdots, x_N') \prod_{n=1}^{N} {}^{(n)}\langle x_n \,|x_n'\,\rangle^{(n)} \\
&= \int_{-\infty}^{\infty} dx_1' \int_{-\infty}^{\infty} dx_2' \cdots \int_{-\infty}^{\infty} dx_N' \qquad \psi(x_1', x_2', \cdots, x_N') \prod_{n=1}^{N} \delta(x_n - x_n') \\
&= \psi(x_1, x_2, \cdots, x_N) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (8.52)
\end{aligned}
$$

Lecture 22:
Multiparticle Systems:
Indistinguishable Particles

Revision Date: 2007/11/21

Indistinguishable Particles

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.

But, 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? Well, our construction of the position or momentum operator for a particular particle in the direct product space is now nonsense. Recall, we defined

$$X_1^{(1)\otimes(2)} \equiv X_1^{(1)} \otimes I^{(2)} \tag{8.53}$$

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.

Indistinguishable Particles (cont.)

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 V_1 and 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 V_1, there is a matching state $|v\rangle^{(2)}$ in V_2 and vice versa. This is not necessarily true for distinguishable particles or degrees of freedom.

Now, we need to make a digression on operators for indistinguishable particles.

Operators for Indistinguishable Particles

Given that V_1 and V_2 are identical, we recognize that for any operator $\Omega^{(1)}$ acting in V_1, there is a matching operator $\Omega^{(2)}$ acting in V_2 that acts in an identical manner on V_2 as $\Omega^{(1)}$ acts on V_1. What do we mean by "identical"? An operator is a rule assigning to any input ket an output ket. $\Omega^{(1)}|v\rangle^{(1)}$ is another ket $|w\rangle^{(1)}$ in V_1. By saying that V_1 and 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 V_2. When we say that the action of $\Omega^{(2)}$ in V_2 is identical to that of $\Omega^{(1)}$ in V_1, we are saying that, as one would expect, $|w\rangle^{(2)} = \Omega^{(2)}|v\rangle^{(2)}$ because $|w\rangle^{(1)} = \Omega^{(1)}|v\rangle^{(1)}$.

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

$$\alpha\,|\widehat{\omega}, \widetilde{\omega}\rangle^{(1)\otimes(2)} + \beta\,|\widetilde{\omega}, \widehat{\omega}\rangle^{(1)\otimes(2)} \tag{8.54}$$

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

$$\alpha \, |\widetilde{\omega}, \widehat{\omega}\rangle^{(1)\otimes(2)} + \beta \, |\widehat{\omega}, \widetilde{\omega}\rangle^{(1)\otimes(2)} = \gamma \left[\alpha \, |\widehat{\omega}, \widetilde{\omega}\rangle^{(1)\otimes(2)} + \beta \, |\widetilde{\omega}, \widehat{\omega}\rangle^{(1)\otimes(2)} \right] \qquad (8.55)$$

We thus obtain the conditions

$$\alpha = \gamma \, \beta \qquad \beta = \gamma \, \alpha \qquad \implies \qquad \gamma^2 = 1 \qquad \alpha^2 = \beta^2 \qquad (8.56)$$

Note that we have the square, not the modulus squared, of each variable, which implies that the solutions are

$$\gamma = 1 \quad \beta = \alpha \qquad \text{and} \qquad \gamma = -1 \quad \beta = -\alpha \qquad (8.57)$$

So the allowed states are (up to normalization)

$$|\widehat{\omega}, \widetilde{\omega}\rangle_{+}^{(1)\otimes(2)} = |\widehat{\omega}, \widetilde{\omega}\rangle_{S}^{(1)\otimes(2)} = |\widehat{\omega}, \widetilde{\omega}\rangle^{(1)\otimes(2)} + |\widetilde{\omega}, \widehat{\omega}\rangle^{(1)\otimes(2)} \qquad (8.58)$$

$$|\widehat{\omega}, \widetilde{\omega}\rangle_{-}^{(1)\otimes(2)} = |\widehat{\omega}, \widetilde{\omega}\rangle_{A}^{(1)\otimes(2)} = |\widehat{\omega}, \widetilde{\omega}\rangle^{(1)\otimes(2)} - |\widetilde{\omega}, \widehat{\omega}\rangle^{(1)\otimes(2)} \qquad (8.59)$$

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.

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:

$$\frac{1}{2}\left(|\widehat{\omega}, \widetilde{\omega}\rangle_+^{(1)\otimes(2)} + |\widehat{\omega}, \widetilde{\omega}\rangle_-^{(1)\otimes(2)}\right) = |\widehat{\omega}, \widetilde{\omega}\rangle^{(1)\otimes(2)} \tag{8.60}$$

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.

Indistinguishable Particles (cont.)

And, 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:

$$\Omega_1^{(1)\otimes(2)}|\widehat{\omega},\widetilde{\omega}\rangle_\pm^{(1)\otimes(2)} = \left(\Omega_1^{(1)} \otimes I^{(2)}\right)|\widehat{\omega},\widetilde{\omega}\rangle_\pm^{(1)\otimes(2)} \tag{8.61}$$

$$= \left(\Omega_1^{(1)} \otimes I^{(2)}\right)\left(|\widehat{\omega},\widetilde{\omega}\rangle^{(1)\otimes(2)} \pm |\widetilde{\omega},\widehat{\omega}\rangle^{(1)\otimes(2)}\right) \tag{8.62}$$

$$= \widehat{\omega}\,|\widehat{\omega},\widetilde{\omega}\rangle^{(1)\otimes(2)} \pm \widetilde{\omega}\,|\widetilde{\omega},\widehat{\omega}\rangle^{(1)\otimes(2)} \tag{8.63}$$

$$= \frac{\widehat{\omega}}{2}\left(|\widehat{\omega},\widetilde{\omega}\rangle_+^{(1)\otimes(2)} + |\widehat{\omega},\widetilde{\omega}\rangle_-^{(1)\otimes(2)}\right) \tag{8.64}$$

$$\pm \frac{\widetilde{\omega}}{2}\left(|\widehat{\omega},\widetilde{\omega}\rangle_+^{(1)\otimes(2)} - |\widehat{\omega},\widetilde{\omega}\rangle_-^{(1)\otimes(2)}\right) \tag{8.65}$$

$$= \frac{\widetilde{\omega}\pm\widehat{\omega}}{2}\,|\widehat{\omega},\widetilde{\omega}\rangle_+^{(1)\otimes(2)} + \frac{\widetilde{\omega}\mp\widehat{\omega}}{2}\,|\widehat{\omega},\widetilde{\omega}\rangle_-^{(1)\otimes(2)} \tag{8.66}$$

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:

$$\overline{\Omega}^{(1)\otimes(2)} = \frac{1}{2}\left(\Omega_1^{(1)\otimes(2)} + \Omega_2^{(1)\otimes(2)}\right) = \frac{1}{2}\left(\Omega^{(1)} \otimes I^{(2)} + I^{(1)} \otimes \Omega^{(2)}\right) \qquad (8.67)$$

$$\delta\Omega^{(1)\otimes(2)} = \frac{1}{2}\left|\Omega_1^{(1)\otimes(2)} - \Omega_2^{(1)\otimes(2)}\right| = \frac{1}{2}\left|\Omega^{(1)} \otimes I^{(2)} - I^{(1)} \otimes \Omega^{(2)}\right| \qquad (8.68)$$

Essentially, we construct the two linear combinations of $\Omega_1^{(1)\otimes(2)}$ and $\Omega_2^{(1)\otimes(2)}$ that 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.

Indistinguishable Particles (cont.)

The action of the above operators on our allowed states is then

$$\overline{\Omega}^{(1)\otimes(2)}|\widehat{\omega},\widetilde{\omega}\,\rangle_{\pm}^{(1)\otimes(2)} = \frac{1}{2}\left(\Omega^{(1)}\otimes I^{(2)} + I^{(1)}\otimes\Omega^{(2)}\right)\left(|\widehat{\omega},\widetilde{\omega}\,\rangle^{(1)\otimes(2)} \pm |\widetilde{\omega},\widehat{\omega}\,\rangle^{(1)\otimes(2)}\right)$$

$$\tag{8.69}$$

$$= \frac{\widehat{\omega}+\widetilde{\omega}}{2}\,|\widehat{\omega},\widetilde{\omega}\,\rangle^{(1)\otimes(2)} \pm \frac{\widetilde{\omega}+\widehat{\omega}}{2}\,|\widetilde{\omega},\widehat{\omega}\,\rangle^{(1)\otimes(2)} \tag{8.70}$$

$$= \frac{\widehat{\omega}+\widetilde{\omega}}{2}\,|\widehat{\omega},\widetilde{\omega}\,\rangle_{\pm}^{(1)\otimes(2)} \tag{8.71}$$

$$\delta\Omega^{(1)\otimes(2)}|\widehat{\omega},\widetilde{\omega}\,\rangle_{\pm}^{(1)\otimes(2)} = \frac{1}{2}\left|\Omega^{(1)}\otimes I^{(2)} - I^{(1)}\otimes\Omega^{(2)}\right|\left(|\widehat{\omega},\widetilde{\omega}\,\rangle^{(1)\otimes(2)} \pm |\widetilde{\omega},\widehat{\omega}\,\rangle^{(1)\otimes(2)}\right)$$

$$\tag{8.72}$$

$$= \frac{|\widehat{\omega}-\widetilde{\omega}|}{2}\,|\widehat{\omega},\widetilde{\omega}\,\rangle^{(1)\otimes(2)} \pm \frac{|\widetilde{\omega}-\widehat{\omega}|}{2}\,|\widetilde{\omega},\widehat{\omega}\,\rangle^{(1)\otimes(2)} \tag{8.73}$$

$$= \frac{|\widehat{\omega}-\widetilde{\omega}|}{2}\,|\widehat{\omega},\widetilde{\omega}\,\rangle_{\pm}^{(1)\otimes(2)} \tag{8.74}$$

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 $|\widetilde{\omega},\widehat{\omega}\,\rangle_{\pm}^{(1)\otimes(2)}$.

Indistinguishable Particles (cont.)

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:

$$V_{(1)\otimes(2)} = V_+ \oplus V_- = V_S \oplus V_A \tag{8.75}$$

Recall that the above does *not* imply that $V_{(1)\otimes(2)}$ is the union of the two subspaces: the direct sum spaces is bigger than either of the subspaces V_+ and 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 V_+ or V_- significantly reduces the number of states available – it is not just smaller by half, but by a dimension, so to speak. For Hilbert spaces that have countably or uncountably infinite dimension, this reductions means little, but the idea is the same.

An analogy would be think of $V_{(1)\otimes(2)}$ as the real plane, V_+ as all the vectors along the line $y = x$, and V_- as all the vectors along the line $y = -x$. The two spaces V_+ and V_- are subspaces and their linear combinations give the entire space $V_{(1)\otimes(2)}$, but their union would just be the points satisfying $y = \pm x$.

One might be tempted to think that V_+ and V_- are in some sense identical, that for every state in V_+ there is a state in 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

$$|\omega, \omega\rangle_+^{(1)\otimes(2)} = |\omega, \omega\rangle^{(1)\otimes(2)} \qquad (8.76)$$

$$|\omega, \omega\rangle_-^{(1)\otimes(2)} = 0 \qquad (8.77)$$

(where 0 means the null vector here). That is, there is a subset of states in V_+ for which there is no partner in 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 boson since such states are automatically symmetric under particle exchange.

Expansions in Terms of Symmetric and Antisymmetric Bases

In constructing states in terms of V_+ or 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 8.13)

$$|u\rangle^{(1)\otimes(2)} = \sum_{i,j} u_{ij}|v_i, w_j\rangle^{(1)\otimes(2)} \tag{8.78}$$

where there is no restriction on i and j. If we specialize to a system with two particles that are identical but distinguishable, the above becomes

$$|u\rangle^{(1)\otimes(2)} = \sum_{i,j} u_{ij}|\omega_i, \omega_j\rangle^{(1)\otimes(2)} \tag{8.79}$$

where we have labeled the single particle states by ω_i and ω_j instead of v_i and w_j because the two factor spaces are now identical (*i.e.*, we should use v_j instead of w_j) and for consistency with our early notation (use ω_i, ω_j instead of v_i, v_j). There is no restriction on i and j yet.

Indistinguishable Particles (cont.)

Let's now rewrite in terms of the symmetric and antisymmetric bases, still treating the particles as distinguishable:

$$
\begin{aligned}
|u\rangle^{(1)\otimes(2)} &= \sum_i u_{ii} |\omega_i, \omega_i\rangle^{(1)\otimes(2)} + \sum_{i>j} \Big[\frac{u_{ij}}{\sqrt{2}} \left(|\omega_i, \omega_j\rangle_+^{(1)\otimes(2)} + |\omega_i, \omega_j\rangle_-^{(1)\otimes(2)} \right) \\
&\qquad + \frac{u_{ji}}{\sqrt{2}} \left(|\omega_i, \omega_j\rangle_+^{(1)\otimes(2)} - |\omega_i, \omega_j\rangle_-^{(1)\otimes(2)} \right) \Big] \\
&= \sum_i u_{ii} |\omega_i, \omega_i\rangle^{(1)\otimes(2)} \\
&\qquad + \sum_{i>j} \left(\frac{u_{ij} + u_{ji}}{\sqrt{2}} |\omega_i, \omega_j\rangle_+^{(1)\otimes(2)} + \frac{u_{ij} - u_{ji}}{\sqrt{2}} |\omega_i, \omega_j\rangle_-^{(1)\otimes(2)} \right) \\
&\equiv \sum_i u_{ii} |\omega_i, \omega_i\rangle^{(1)\otimes(2)} + \sum_{i>j} \left(u_{ij,+} |\omega_i, \omega_j\rangle_+^{(1)\otimes(2)} + u_{ij,-} |\omega_i, \omega_j\rangle_-^{(1)\otimes(2)} \right)
\end{aligned}
\tag{8.80}
$$

with, for $i \neq j$

$$
u_{ij,+} = \frac{u_{ij} + u_{ji}}{\sqrt{2}} \quad |\omega_i, \omega_j\rangle_+^{(1)\otimes(2)} = \frac{1}{\sqrt{2}} \left(|\omega_i, \omega_j\rangle^{(1)\otimes(2)} + |\omega_j, \omega_i\rangle^{(1)\otimes(2)} \right)
\tag{8.81}
$$

$$
u_{ij,-} = \frac{u_{ij} - u_{ji}}{\sqrt{2}} \quad |\omega_i, \omega_j\rangle_-^{(1)\otimes(2)} = \frac{1}{\sqrt{2}} \left(|\omega_i, \omega_j\rangle^{(1)\otimes(2)} - |\omega_j, \omega_i\rangle^{(1)\otimes(2)} \right)
\tag{8.82}
$$

Indistinguishable Particles (cont.)

We have pulled out the $i = j$ 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 $V_{(1)\otimes(2)} = V_+ \oplus V_-$. Also, now instead of coefficients u_{ij} with $i > j$ and $i < j$ allowed, we consider only the coefficients $u_{ij,+}$ and $u_{ij,-}$ with $i > j$ 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 $u_{ij,-} = 0$, and for antisymmetric states, we set all $u_{ii} = 0$ and all $u_{ij,+} = 0$. The resulting two kinds of states are

$$|u\rangle_+^{(1)\otimes(2)} = \sum_i u_{ii} |\omega_i, \omega_i\rangle^{(1)\otimes(2)} + \sum_{i>j} u_{ij,+} |\omega_i, \omega_j\rangle_+^{(1)\otimes(2)} \qquad (8.83)$$

$$|u\rangle_-^{(1)\otimes(2)} = \sum_{i>j} u_{ij,-} |\omega_i, \omega_j\rangle_-^{(1)\otimes(2)}$$

If the $|\omega_i, \omega_j\rangle^{(1)\otimes(2)}$ have unity norm, then $|\omega_i, \omega_i\rangle^{(1)\otimes(2)}$ and $|\omega_i, \omega_j\rangle_\pm^{(1)\otimes(2)}$ have unity norm, so the normalization condition on the states is

$$\sum_i |u_{ii}|^2 + \sum_{i>j} |u_{ij,+}|^2 = 1 \qquad \text{or} \qquad \sum_{i>j} |u_{ij,-}|^2 = 1 \qquad (8.84)$$

So, the algorithm for normalizing the states is clear.

Indistinguishable Particles (cont.)

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:

$$\langle \overline{\Omega}^{(1)\otimes(2)} \rangle_+ = {}^{(1)\otimes(2)}_+\langle u | \overline{\Omega}^{(1)\otimes(2)} | u \rangle_+^{(1)\otimes(2)} \tag{8.85}$$

$$= \sum_i |u_{ii}|^2 \, {}^{(1)\otimes(2)}_+\langle \omega_i, \omega_i | \overline{\Omega}^{(1)\otimes(2)} | \omega_i, \omega_i \rangle_+^{(1)\otimes(2)} \tag{8.86}$$

$$+ \sum_{i>j} |u_{ij,+}|^2 \, {}^{(1)\otimes(2)}_+\langle \omega_i, \omega_j | \overline{\Omega}^{(1)\otimes(2)} | \omega_i, \omega_j \rangle_+^{(1)\otimes(2)}$$

Recall that $|\omega_i, \omega_j \rangle_\pm^{(1)\otimes(2)}$ is an eigenvector of $\overline{\Omega}$ with eigenvalue $(\omega_i + \omega_j)/2$. So the sum becomes easy:

$$\langle \overline{\Omega}^{(1)\otimes(2)} \rangle_+ = \sum_i |u_{ii}|^2 \omega_i + \sum_{i>j} |u_{ij,+}|^2 \left(\frac{\omega_i + \omega_j}{2} \right) \tag{8.87}$$

For an antisymmetric state, we have

$$\langle \overline{\Omega}^{(1) \otimes (2)} \rangle_- = \sum_{i>j} |u_{ij,-}|^2 \, {}^{(1) \otimes (2)}_- \langle \omega_i, \omega_j | \overline{\Omega}^{(1) \otimes (2)} | \omega_i, \omega_j \rangle^{(1) \otimes (2)}_- \tag{8.88}$$

$$= \sum_{i>j} |u_{ij,-}|^2 \left(\frac{\omega_i + \omega_j}{2} \right) \tag{8.89}$$

For the difference operator, we have

$$\langle \delta\Omega^{(1) \otimes (2)} \rangle_+ = \sum_i |u_{ii}|^2 (0) + \sum_{i>j} |u_{ij,+}|^2 \left| \frac{\omega_i - \omega_j}{2} \right| = \sum_{i>j} |u_{ij,+}|^2 \left| \frac{\omega_i - \omega_j}{2} \right| \tag{8.90}$$

$$\langle \delta\Omega^{(1) \otimes (2)} \rangle_- = \sum_{i>j} |u_{ij,-}|^2 \left| \frac{\omega_i - \omega_j}{2} \right| \tag{8.91}$$

Indistinguishable Particles (cont.)

It is sometimes more convenient to let both i and j 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 Equations 8.86 as

$$\langle \overline{\Omega}^{(1)\otimes(2)} \rangle_+ = \sum_i |u_{ii}|^2 \, {}^{(1)\otimes(2)}_+ \langle \omega_i, \omega_i | \, \overline{\Omega}^{(1)\otimes(2)} | \omega_i, \omega_i \rangle^{(1)\otimes(2)}_+ \tag{8.92}$$

$$+ \sum_{i>j} |u_{ij,+}|^2 \, {}^{(1)\otimes(2)}_+ \langle \omega_i, \omega_j | \, \overline{\Omega}^{(1)\otimes(2)} | \omega_i, \omega_j \rangle^{(1)\otimes(2)}_+ \tag{8.93}$$

$$= \frac{1}{2} \sum_i |u_{ii,+}|^2 \, {}^{(1)\otimes(2)}_+ \langle \omega_i, \omega_i | \, \overline{\Omega}^{(1)\otimes(2)} | \omega_i, \omega_i \rangle^{(1)\otimes(2)}_+ \tag{8.94}$$

$$+ \frac{1}{2} \sum_{i\neq j} |u_{ij,+}|^2 \, {}^{(1)\otimes(2)}_+ \langle \omega_i, \omega_j | \, \overline{\Omega}^{(1)\otimes(2)} | \omega_i, \omega_j \rangle^{(1)\otimes(2)}_+ \tag{8.95}$$

$$= \frac{1}{2} \sum_{i,j} |u_{ij,+}|^2 \, {}^{(1)\otimes(2)}_+ \langle \omega_i, \omega_j | \, \overline{\Omega}^{(1)\otimes(2)} | \omega_i, \omega_j \rangle^{(1)\otimes(2)}_+ \tag{8.96}$$

where we have defined $u_{ii,+}$ by extending the formula for $u_{ij,+}$ without modification, $u_{ii,+} = \frac{u_{ii} + u_{ii}}{\sqrt{2}} = \sqrt{2}\, u_{ii}$.

Indistinguishable Particles (cont.)

and 8.89 as

$$\langle\overline{\Omega}^{(1)\otimes(2)}\rangle_- = \sum_{i>j} |u_{ij,-}|^2 \, {}^{(1)\otimes(2)}_-\langle\omega_i,\omega_j\,|\,\overline{\Omega}^{(1)\otimes(2)}|\omega_i,\omega_j\,\rangle^{(1)\otimes(2)}_- \qquad (8.97)$$

$$= \frac{1}{2}\sum_{i\neq j} |u_{ij,-}|^2 \, {}^{(1)\otimes(2)}_-\langle\omega_i,\omega_j\,|\,\overline{\Omega}^{(1)\otimes(2)}|\omega_i,\omega_j\,\rangle^{(1)\otimes(2)}_- \qquad (8.98)$$

$$= \frac{1}{2}\sum_{i,j} |u_{ij,-}|^2 \, {}^{(1)\otimes(2)}_+\langle\omega_i,\omega_j\,|\,\overline{\Omega}^{(1)\otimes(2)}|\omega_i,\omega_j\,\rangle^{(1)\otimes(2)}_- \qquad (8.99)$$

where $u_{ii,-} = 0$ follows from the definition of $u_{ij,-}$, $u_{ii,-} = \frac{u_{ii}-u_{ii}}{\sqrt{2}} = 0$.

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

$$\langle\overline{\Omega}^{(1)\otimes(2)}\rangle_+ = \frac{1}{2}\int d\widehat{\omega}\int d\widetilde{\omega}\,|u_+(\widehat{\omega},\widetilde{\omega})|^2 \, {}^{(1)\otimes(2)}_+\langle\widehat{\omega},\widetilde{\omega}\,|\,\overline{\Omega}^{(1)\otimes(2)}|\widehat{\omega},\widetilde{\omega}\,\rangle^{(1)\otimes(2)}_+ \quad (8.100)$$

$$\langle\overline{\Omega}^{(1)\otimes(2)}\rangle_- = \frac{1}{2}\int d\widehat{\omega}\int d\widetilde{\omega}\,|u_-(\widehat{\omega},\widetilde{\omega})|^2 \, {}^{(1)\otimes(2)}_-\langle\widehat{\omega},\widetilde{\omega}\,|\,\overline{\Omega}^{(1)\otimes(2)}|\widehat{\omega},\widetilde{\omega}\,\rangle^{(1)\otimes(2)}_- \quad (8.101)$$

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 $\delta\Omega$, we will obtain values $\frac{\omega_i + \omega_j}{2}$ and $\frac{|\omega_i - \omega_j|}{2}$ ($\frac{\widehat{\omega} + \widetilde{\omega}}{2}$ and $\frac{|\widehat{\omega} - \widetilde{\omega}|}{2}$ for the continuous case). (Note that $\overline{\Omega}$ and $\delta\Omega$ commute and are hence simultaneously diagonalizable.) From these, we may calculate the two numbers ω_i and ω_j (or $\widehat{\omega}$ and $\widetilde{\omega}$), though, we don't know which is which; that is, we obtain

$$\omega_i \quad \text{or} \quad \widehat{\omega} = \langle\overline{\Omega}\rangle \pm \langle\delta\Omega\rangle \qquad \omega_j \quad \text{or} \quad \widetilde{\omega} = \langle\overline{\Omega}\rangle \mp \langle\delta\Omega\rangle \tag{8.102}$$

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 8.84) and for the expectation values $\langle\overline{\Omega}\rangle_\pm$ and $\langle\delta\Omega\rangle_\pm$ make this clear: the probability is

$$P_\pm(\omega_i, \omega_j) = |u_{ij,\pm}|^2 \qquad \text{or} \qquad P_\pm(\widehat{\omega}, \widetilde{\omega}) = |u_\pm(\widehat{\omega}, \widetilde{\omega})|^2 \tag{8.103}$$

We are assured that the probability, summed or integrated over all possible outcomes, is normalized to unity by Equation 8.84.

Lecture 23:
Multiparticle Systems:
Indistinguishable Particles continued

Revision Date: 2007/11/21

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

$$P_\pm(\widehat{x}, \widetilde{x}) = |u_\pm(\widehat{x}, \widetilde{x})|^2 \tag{8.104}$$

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,

$$u_\pm(\widehat{x}, \widetilde{x}) = {}_\pm\langle \widehat{x}, \widetilde{x} | \psi_a, \psi_b \rangle_\pm \tag{8.105}$$

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!):

$$\pm \langle \widehat{x}, \widetilde{x} \,|\, \psi_a, \psi_b \rangle_\pm = \frac{1}{\sqrt{2}} \left(\langle \widehat{x}, \widetilde{x} \,| \pm \langle \widetilde{x}, \widehat{x} \,| \right) \frac{1}{\sqrt{2}} \left(|\psi_a, \psi_b \rangle \pm |\psi_b, \psi_a \rangle \right) \qquad (8.106)$$

$$= \frac{1}{2} \Big[\psi_{a,x}(\widehat{x}) \, \psi_{b,x}(\widetilde{x}) + \psi_{b,x}(\widetilde{x}) \psi_{a,x}(\widehat{x}) \qquad\qquad (8.107)$$

$$\pm \psi_{b,x}(\widehat{x}) \, \psi_{a,x}(\widetilde{x}) \pm \psi_{a,x}(\widetilde{x}) \, \psi_{b,x}(\widehat{x}) \Big]$$

$$= \psi_{a,x}(\widehat{x}) \, \psi_{b,x}(\widetilde{x}) \pm \psi_{b,x}(\widehat{x}) \, \psi_{a,x}(\widetilde{x}) \qquad\qquad (8.108)$$

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:

$$P_\pm(\widehat{x}, \widetilde{x}) = |u_\pm(\widehat{x}, \widetilde{x})|^2 = \left[\psi_{a,x}(\widehat{x}) \, \psi_{b,x}(\widetilde{x}) \pm \psi_{b,x}(\widehat{x}) \, \psi_{a,x}(\widetilde{x}) \right]^* \qquad (8.109)$$

$$\times \left[\psi_{a,x}(\widehat{x}) \, \psi_{b,x}(\widetilde{x}) \pm \psi_{b,x}(\widehat{x}) \, \psi_{a,x}(\widetilde{x}) \right]$$

$$= |\psi_{a,x}(\widehat{x})|^2 |\psi_{b,x}(\widetilde{x})|^2 + |\psi_{b,x}(\widehat{x})|^2 |\psi_{a,x}(\widetilde{x})|^2 \qquad (8.110)$$

$$\pm \psi_{a,x}^*(\widehat{x}) \, \psi_{b,x}(\widehat{x}) \, \psi_{b,x}^*(\widetilde{x}) \, \psi_{a,x}(\widetilde{x})$$

$$\pm \psi_{b,x}^*(\widehat{x}) \, \psi_{a,x}(\widehat{x}) \, \psi_{a,x}^*(\widetilde{x}) \, \psi_{b,x}(\widetilde{x})$$

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:

$$P_{\pm}(x_0, x_0) = 2 \left(|\psi_{a,x}(x_0)|^2 |\psi_{b,x}(x_0)|^2 \pm |\psi_{a,x}(x_0)|^2 |\psi_{b,x}(x_0)|^2 \right) \tag{8.111}$$

$$= \begin{cases} 4 \, |\psi_{a,x}(x_0)|^2 |\psi_{b,x}(x_0)|^2 & + \\ 0 & - \end{cases} \tag{8.112}$$

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!

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.)

Symmetrization and Antisymmetrizaton 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 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 *i*th of $N!$ possible permutations as P_i, where $k = P_i(j)$ indicates that particle j in the unpermuted state should be replaced by particle k in the permuted state, and the permutations are defined so that P_i requires an odd number of exchanges when i is odd and an even number when i is even, then we have

$$|\psi_{n_1}, \cdots, \psi_{n_N}\rangle_\pm = \frac{1}{\sqrt{N!}} \sum_{i=1}^{N!} (\pm 1)^i |\psi_{n_1}\rangle^{(P_i(1))} \otimes \cdots \otimes |\psi_{n_N}\rangle^{(P_i(N))} \qquad (8.113)$$

$$= \frac{1}{\sqrt{N!}} \sum_{i=1}^{N!} (\pm 1)^i \prod_{j=1}^{N} |\psi_{n_j}\rangle^{(P_i(j))} \qquad (8.114)$$

where, in order to provide compact notation, the states are allowed to go out of particle order; the $^{(P_i(j))}$ indicates which particle the state n_j is assigned to in the *i* the permutation.

Indistinguishable Particles (cont.)

The above notation works only if all the states are distinct (if $n_j \neq n_k$ for all j and k) 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, $n_j = n_k$ for some j, k pair, then, for every permutation i, there will be a matching permutation that differs only by exchanging the two particles $P_i(j)$ and $P_i(k)$ to which the states $n_j = n_k$ have been assigned. The sum collapses 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 $\sqrt{\frac{2}{N!}}$.

For fermions, the antisymmetrized state can be written conveniently as a determinant, usually termed a Slater Determinant:

$$|\psi_{n_1}, \cdots, \psi_{n_N}\rangle_- = \frac{1}{\sqrt{N!}} \begin{vmatrix} |\psi_{n_1}\rangle^{(1)} & \cdots & |\psi_{n_N}\rangle^{(1)} \\ \vdots & \ddots & \vdots \\ |\psi_{n_1}\rangle^{(N)} & \cdots & |\psi_{n_N}\rangle^{(N)} \end{vmatrix} \tag{8.115}$$

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 appropriate result.

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

$$|\psi_a, \psi_b\rangle_{\pm}^{(1)\otimes(2)} = \frac{1}{\sqrt{2}} \left[|\psi_a\rangle^{(1)} \otimes |\psi_b\rangle^{(2)} \pm |\psi_b\rangle^{(1)} \otimes |\psi_a\rangle^{(2)} \right] \qquad (8.116)$$

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 8.110,

$$\begin{aligned} P_{\pm}(x_{SM}, x_{OC}) = & |\psi_{a,x}(x_{SM})|^2 |\psi_{b,x}(x_{OC})|^2 + |\psi_{b,x}(x_{SM})|^2 |\psi_{a,x}(x_{OC})|^2 \qquad (8.117) \\ & \pm \psi_{a,x}^*(x_{SM})\,\psi_{b,x}(x_{SM})\,\psi_{b,x}^*(x_{OC})\,\psi_{a,x}(x_{OC}) \\ & \pm \psi_{b,x}^*(x_{SM})\,\psi_{a,x}(x_{SM})\,\psi_{a,x}^*(x_{OC})\,\psi_{b,x}(x_{OC}) \end{aligned}$$

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:

$$
\begin{aligned}
P_{\pm}(x_{SM}) = {} & |\psi_{a,x}(x_{SM})|^2 \int_{-\infty}^{\infty} dx_{OC} \, |\psi_{b,x}(x_{OC})|^2 \qquad (8.118) \\
& + |\psi_{b,x}(x_{SM})|^2 \int_{-\infty}^{\infty} dx_{OC} \, |\psi_{a,x}(x_{OC})|^2 \\
& \pm \psi_{a,x}^*(x_{SM}) \, \psi_{b,x}(x_{SM}) \int_{-\infty}^{\infty} dx_{OC} \, \psi_{b,x}^*(x_{OC}) \, \psi_{a,x}(x_{OC}) \\
& \pm \psi_{b,x}^*(x_{SM}) \, \psi_{a,x}(x_{SM}) \int_{-\infty}^{\infty} dx_{OC} \, \psi_{a,x}^*(x_{OC}) \, \psi_{b,x}(x_{OC})
\end{aligned}
$$

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.

So we have

$$P_{\pm}(x_{SM}) = |\psi_{a,x}(x_{SM})|^2 \tag{8.119}$$

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 9
Symmetries

Lecture 24:
Coordinate Transformations
Symmetry Transformations

Revision Date: 2007/11/30

Coordinate Transformations

Coordinate Transformations

A coordinate transformation is any simple 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:

$$x' = -x \qquad\qquad y' = y \qquad\qquad z' = z \qquad (9.1)$$

- Rotation transformation about the z axis by θ (CCW):

$$x' = x\cos\theta + y\sin\theta \qquad y' = -x\sin\theta + y\cos\theta \qquad z' = z \qquad (9.2)$$

- Translation:

$$x' = x - a \qquad\qquad y' = y \qquad\qquad z' = z \qquad (9.3)$$

Coordinate Transformations (cont.)

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.

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, it affects the operators. With the new coordinate system comes a new set of operators $\{Q'\}$ and $\{P'_q\}$ where $\{Q\}$ stands for all the coordinate dimensions and similarly for $\{P_q\}$. (All other operators are derived from $\{Q\}$ and $\{P_q\}$, so we need only to explicitly discuss these two.) 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.

- Second, because of the above, the position and momentum basis change; there are new sets of basis states $\{|q'\rangle\}$ and $\{|p'_q\rangle\}$ that are eigenvectors of the $\{Q'\}$ and $\{P'_q\}$ operators. 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.

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!

Coordinate Transformations (cont.)

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:

$$e^{i\theta(q)}|q'\rangle = T_P|q\rangle \tag{9.4}$$

where $\theta(q)$ is real. (We will define an *active* transformation below.) The $e^{i\theta(q)}$ 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 $\theta(q) = 0$ 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 there are no non-unity-modulus coefficients in front. Therefore, it preserves all inner products, and hence is unitary, $T^\dagger = T^{-1}$.

Coordinate Transformations (cont.)

The new operators are defined 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,

$$\langle q_1' \, |Q'|q_2' \, \rangle = \langle q_1 \, |Q|q_2 \, \rangle \tag{9.5}$$

By the definition of the transformation operator, we also have

$$\langle q_1' \, |Q'|q_2' \, \rangle = \langle q_1 \, |T_P^\dagger Q' \, T|q_2 \, \rangle \tag{9.6}$$

Combining the two statements gives

$$\langle q_1 \, |Q|q_2 \, \rangle = \langle q_1 \, |T_P^\dagger Q' \, T|q_2 \, \rangle \tag{9.7}$$

Since this relation holds for all $|q_1 \, \rangle$ and $|q_2 \, \rangle$, it therefore holds that

$$Q' = T_P \, Q \, T_P^\dagger = T_P \, Q \, T_P^{-1} \tag{9.8}$$

The above proof carries through for the $\{P_q\}$ also, and, in fact, for any operator.

If the abstractness of Equation 9.8 is confusing, always think of in terms of Equation 9.7 for the matrix elements.

We note that, because we do not want to move the potentials and particles, we do not in general want to transform H to H'. We can use the above relations, though, to write H in terms of the $\{Q'\}$ and $\{P'_q\}$ operators by converting the $\{Q\}$ and $\{P_q\}$ to their primed versions.

Coordinate Transformations (cont.)

Example 9.1: 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

$$E_{ab} = \frac{\hbar^2 \pi^2}{2\,m} \left(\frac{a^2}{L_1^2} + \frac{b^2}{L_2^2} \right) \tag{9.9}$$

The position-basis representation is

$$\psi_{ab}(x, y) = \langle x, y \,|\psi_{ab}\rangle = \sqrt{\frac{4}{L_1 L_2}} \, \sin\left(\frac{a\,\pi\,x}{L_1} \right) \sin\left(\frac{b\,\pi\,y}{L_2} \right) \tag{9.10}$$

(Note that, because the box's corner, not center, is at the origin, the sines are the only allowed states.) and we would find the expectation values

$$\langle \psi_{ab} \,|X\,|\psi_{ab}\rangle = \frac{L_1}{2} \quad \langle \psi_{ab} \,|Y\,|\psi_{ab}\rangle = \frac{L_2}{2} \quad \langle \psi_{ab} \,|P_x\,|\psi_{ab}\rangle = 0 \quad \langle \psi_{ab} \,|P_y\,|\psi_{ab}\rangle = 0 \tag{9.11}$$

Now, perform the coordinate transformation $x' = y$, $y' = x$. The transformation gives $|x' = u, y' = v\rangle = T_P|x = u, y = v\rangle$. Note the ordering of u and v; for example, a state along the x axis, $(x = 1, y = 0)$, gets mapped to a state along the x' axis, $(x' = 1, y' = 0)$. 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 along the x' axis at $(x' = 1, y' = 0)$ is along the y axis at $(x = 0, y = 1)$. Make sure you have the above straight; it *is* confusing!

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.

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:

$$\psi_{ab,q'}(x' = u, y' = v) = \langle x' = u, y' = v \,|\, \psi_{ab}\rangle \tag{9.12}$$

$$= \langle x = v, y = u \,|\, \psi_{ab}\rangle \tag{9.13}$$

$$= \psi_{ab,q}(x = v, y = u) \tag{9.14}$$

$$= \sqrt{\frac{4}{L_1 L_2}} \, \sin\left(\frac{a\,\pi\,v}{L_1}\right) \sin\left(\frac{b\,\pi\,u}{L_2}\right) \tag{9.15}$$

So, $\quad \psi_{ab,q'}(x', y') = \sqrt{\dfrac{4}{L_1 L_2}} \, \sin\left(\dfrac{b\,\pi\,x'}{L_2}\right) \sin\left(\dfrac{a\,\pi\,y'}{L_1}\right) \tag{9.16}$

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.

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

$$\langle x_1', y_1' \, | X' | x_2', y_2' \rangle = \langle x_1, y_1 \, | X | x_2, y_2 \rangle \tag{9.17}$$

Again, it will be far clearer if we use specific numbers not tied to a coordinate system. So let's write the above as

$$\langle x' = u_1, y' = v_1 \, | X' | x' = u_2, y' = v_2 \rangle = \langle x = u_1, y = v_1 \, | X | x_2 = u_2, y_2 = v_2 \rangle \tag{9.18}$$

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

$$\langle x = u_1, y = v_1 \, | X | x = u_2, y = v_2 \rangle = u_1 \, \delta(u_1 - u_2) \, \delta(v_1 - v_2) \tag{9.19}$$

So our requirement becomes

$$\langle x' = u_1, y' = v_1 \, | X' | x' = u_2, y' = v_2 \rangle = u_1 \, \delta(u_1 - u_2) \, \delta(v_1 - v_2) \qquad (9.20)$$

Since u_1 is the x' element of $|x' = u_1, y' = v_1 \rangle$, 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 what we think it 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.

By analogy, we may also conclude

$$\langle x' = u_1, y' = v_1 \,|Y'|x' = u_2, y' = v_2 \rangle = v_1 \, \delta(u_1 - u_2) \, \delta(v_1 - v_2) \tag{9.21}$$

$$\langle x' = u_1, y' = v_1 \,|P_x'|x' = u_2, y' = v_2 \rangle = -i\,\hbar\,\frac{d}{du_1}\,\delta(u_1 - u_2)\,\delta(v_1 - v_2) \tag{9.22}$$

$$\langle x' = u_1, y' = v_1 \,|P_x'|x' = u_2, y' = v_2 \rangle = -i\,\hbar\,\frac{d}{dv_1}\,\delta(u_1 - u_2)\,\delta(v_1 - v_2) \tag{9.23}$$

That is, all the primed operators do exactly what we expect them to in terms of the primed coordinate system. It is ok to take 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.

Coordinate Transformations (cont.)

Let us clarify one point that will become important below when we calculate expectation values. Note the following:

$$\langle x' = u_1, y' = v_1 \, |X'| x' = u_2, y' = v_2 \rangle \tag{9.24}$$
$$= \langle x = u_1, y = v_1 \, |X| x = u_2, y = v_2 \rangle$$
$$= u_1 \, \delta(u_1 - u_2) \, \delta(v_1 - v_2)$$
$$= \langle x = v_1, y = u_1 \, |Y| x = v_2, y = u_2 \rangle$$

Since it is true that

$$|x' = u_1, y' = v_1 \rangle = |x = v_1, y = u_1 \rangle \qquad |x' = u_2, y' = v_2 \rangle = |x = v_2, y = u_2 \rangle \tag{9.25}$$

(note the location of the u's and v's and notice there are no T_P operators in these relations!), we have

$$\langle x' = u_1, y' = v_1 \, |X'| x' = u_2, y' = v_2 \rangle = \langle x' = u_1, y' = v_1 \, |Y| x' = u_2, y' = v_2 \rangle \tag{9.26}$$

That is, all the possible matrix elements between basis elements of X' and Y are equal, so they must be the same operator, $X' = Y$.

But we also have a relation between X' and X. Let's write both of these down:

$$X' = T_P X\, T_P^\dagger \qquad X' = Y \tag{9.27}$$

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':

$$Y' = T_P Y\, T_P^\dagger \qquad\qquad Y' = X \tag{9.28}$$

$$P_x' = T_P P_x T_P^\dagger \qquad\qquad P_x' = P_y \tag{9.29}$$

$$P_y' = T_P P_y T_P^\dagger \qquad\qquad P_y' = P_x \tag{9.30}$$

We note that simple relations like these do not always hold – they are specific to our mirror transformation example. In general, though, it is 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 $\{P_q\}$.

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:

$$H' = T_P \left(\frac{P_x^2 + P_y^2}{2\,m} + V(X, Y) \right) T_P^\dagger = \frac{(P_x')^2 + (P_y')^2}{2\,m} + V(X', Y') \qquad (9.31)$$

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:

$$H = \frac{P_x^2 + P_y^2}{2\,m} + V(X, Y) = \frac{(P_y')^2 + (P_x')^2}{2\,m} + V(Y', X') \qquad (9.32)$$

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'.

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:

$$\langle\psi_{ab}|X'|\psi_{ab}\rangle = \langle\psi_{ab}|Y|\psi_{ab}\rangle = \frac{L_2}{2} \tag{9.33}$$

$$\langle\psi_{ab}|Y'|\psi_{ab}\rangle = \langle\psi_{ab}|X|\psi_{ab}\rangle = \frac{L_1}{2} \tag{9.34}$$

$$\langle\psi_{ab}|P_x'|\psi_{ab}\rangle = \langle\psi_{ab}|P_y|\psi_{ab}\rangle = 0 \tag{9.35}$$

$$\langle\psi_{ab}|P_y'|\psi_{ab}\rangle = \langle\psi_{ab}|P_x|\psi_{ab}\rangle = 0 \tag{9.36}$$

Coordinate Transformations (cont.)

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. The distinction is that the axes themselves don't change – the $\{q\}$ axes stay put relative to the underlying space – but the system moves. The position basis $\{|q\rangle\}$, momentum basis $\{|p_q\rangle\}$, and the $\{Q\}$ and $\{P_q\}$ operators are unchanged. Instead, the state changes; let's 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).

As for T_P, we define T_A explicitly by its action on the position basis elements, via

$$e^{i\,\theta(q)}|q'\rangle = T_A |q\rangle \tag{9.37}$$

where $\theta(q)$ is real, and the prefactor exists for the same reasons discussed in connection to passive transformations.

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 transformed state then satisfies

$$\langle q' | \psi' \rangle = \langle q | T_A^\dagger T_A | \psi \rangle = \langle q | T_A^{-1} T_A | \psi \rangle = \langle q | \psi \rangle \tag{9.38}$$

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.

What about operators? As we mentioned, the original operators $\{Q\}$ and $\{P_Q\}$ are unchanged. But we may create new operators $\{Q'\}$ and $\{P'_q\}$, where we define them by requiring that their matrix elements in the transformed basis are the same as the matrix elements of the old operators in the untransformed basis:

$$\langle q'_1 | Q' | q'_2 \rangle \equiv \langle q_1 | Q | q_2 \rangle \tag{9.39}$$

Rewriting using the relation between the transformed and untransformed basis,

$$\langle q_1 | Q | q_2 \rangle = \langle q'_1 | Q' | q'_2 \rangle = \langle q_1 | T_A^\dagger Q' T_A | q_2 \rangle \tag{9.40}$$

That is, $T_A^\dagger Q' T_A$ has the same matrix elements in the $\{|q\rangle\}$ basis as Q, so we may infer

$$Q' = T_A Q T_A^\dagger = T_A Q T_A^{-1} \tag{9.41}$$

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. Again, if there is any confusion, one only needs to write the above in terms of matrix elements.

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

$$H'\left(T_A|\psi_E\rangle\right) = \left(T_A H\, T_A^\dagger\right)\left(T_A|\psi_E\rangle\right) = T_A H\,|\psi_E\rangle = T_A E\,|\psi_E\rangle = E\left(T_A|\psi_E\rangle\right) \quad (9.42)$$

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:

$$H\,|\psi_E'\rangle = H\,T_A|\psi_E\rangle \qquad H'|\psi_E\rangle = H\,T_A^\dagger|\psi_E'\rangle \quad (9.43)$$

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".

Example 9.2: Active Mirror Transformation of a Particle in a Two-Dimensional Box

In our two-dimensional box example, the active version of the transformation 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 $|\psi_{ab}'\rangle = T_A |\psi_{ab}\rangle$. The projections of the transformed state onto the transformed basis are equal to the projection of the old state onto the untransformed basis:

$$\langle x' = u, y' = v | \psi_{ab}'\rangle = \langle x = u, y = v | T_A^{\dagger} T_A | \psi_{ab}\rangle = \langle x = u, y = v | \psi_{ab}\rangle \quad (9.44)$$

Note that we are not projecting onto $|x = u, y = v\rangle$ 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

$$\psi_{ab,q}(x, y) = \sqrt{\frac{4}{L_1 L_2}} \, \sin\left(\frac{a\pi x}{L_1}\right) \, \sin\left(\frac{b\pi y}{L_2}\right) \tag{9.45}$$

$$\psi_{ab,q'}(x', y') = \sqrt{\frac{4}{L_1 L_2}} \, \sin\left(\frac{b\pi x'}{L_2}\right) \, \sin\left(\frac{a\pi y'}{L_1}\right) \tag{9.46}$$

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

$$\psi'_{ab,q'}(x' = u, y' = v) = \psi_{ab}(x = u, y = v) \tag{9.47}$$

$$= \sqrt{\frac{4}{L_1 L_2}} \, \sin\left(\frac{a\pi u}{L_1}\right) \sin\left(\frac{b\pi v}{L_2}\right) \tag{9.48}$$

$$= \sqrt{\frac{4}{L_1 L_2}} \, \sin\left(\frac{a\pi x'}{L_1}\right) \sin\left(\frac{b\pi y'}{L_2}\right) \tag{9.49}$$

$$\text{or,} \quad \psi'_{ab,q'}(x', y') = \sqrt{\frac{4}{L_1 L_2}} \, \sin\left(\frac{a\pi x'}{L_1}\right) \sin\left(\frac{b\pi y'}{L_2}\right) \tag{9.50}$$

Note how $\psi'_{ab,q'}(x', y') \neq \psi_{ab,q'}(x', y')$! ψ' is really a different state than ψ.

For completeness, one is inclined to consider the two projections $\langle x = u, y = v | \psi'_{ab} \rangle$ and $\langle x' = u, y' = v | \psi_{ab} \rangle$. Though one might think it, they are not in general related to each other:

$$\langle x = u, y = v | \psi'_{ab} \rangle = \langle x = u, y = v | T_A | \psi_{ab} \rangle \tag{9.51}$$

$$\langle x' = u, y' = v | \psi_{ab} \rangle = \langle x = u, y = v | T_A^\dagger | \psi_{ab} \rangle \tag{9.52}$$

Coordinate Transformations (cont.)

However, in this specific case of a mirror transformation, they are related:

$$\langle x = u, y = v \,|\psi'_{ab}\rangle = \langle x' = v, y' = u \,|\psi'_{ab}\rangle \tag{9.53}$$

$$= \sqrt{\frac{4}{L_1 L_2}} \, \sin\left(\frac{a\,\pi\,v}{L_1}\right) \, \sin\left(\frac{b\,\pi\,u}{L_2}\right) \tag{9.54}$$

$$\langle x' = u, y' = v \,|\psi_{ab}\rangle = \langle x = v, y = u \,|\psi_{ab}\rangle \tag{9.55}$$

$$= \sqrt{\frac{4}{L_1 L_2}} \, \sin\left(\frac{a\,\pi\,v}{L_1}\right) \, \sin\left(\frac{b\,\pi\,u}{L_2}\right) \tag{9.56}$$

This is a special case that occurs because $T_A^2 = I$ for our mirror transformation: two mirrors through the same plane are the identity. This tells us $T_A = T_A^{-1}$; unitarity tells us $T_A^{-1} = T_A^\dagger$, so together we have $T_A = T_A^\dagger$, or Hermiticity, which then makes $\langle x = u, y = v \,|T_A|\psi_{ab}\rangle = \langle x = u, y = v \,|T_A^\dagger|\psi_{ab}\rangle$ as needed to make the generally untrue relation true in this case.

The action of active transformations on operators is identical to that for passive transformations, everything we did above carries through. Again, though, because we have the new state $|\psi'_{ab}\rangle$, we can in addition calculate expectation values of the unprimed and primed operators in this state. It is easy to do these calculations just using unitarity, though:

$$\langle \psi'_{ab}|O'|\psi'_{ab}\rangle = \langle \psi_{ab}|\, T_A^\dagger O' T_A|\psi_{ab}\rangle = \langle \psi_{ab}|O|\psi_{ab}\rangle \tag{9.57}$$

This is of course just an extension of the generic fact $\langle q_1|O|q_2\rangle = \langle q'_1|O'|q'_2\rangle$ that we used to define the primed operators (Equation 9.39). So we have

$$\langle \psi'_{ab}|X'|\psi'_{ab}\rangle = \langle \psi_{ab}|X|\psi_{ab}\rangle = \frac{L_1}{2} \tag{9.58}$$

$$\langle \psi'_{ab}|Y'|\psi'_{ab}\rangle = \langle \psi_{ab}|Y|\psi_{ab}\rangle = \frac{L_2}{2} \tag{9.59}$$

$$\langle \psi'_{ab}|P'_x|\psi'_{ab}\rangle = \langle \psi_{ab}|P_x|\psi_{ab}\rangle = 0 \tag{9.60}$$

$$\langle \psi'_{ab}|P'_y|\psi'_{ab}\rangle = \langle \psi_{ab}|P_y|\psi_{ab}\rangle = 0 \tag{9.61}$$

Coordinate Transformations (cont.)

Again, for completness, we are inclined to ask whether there is a relation between $\langle \psi'_{ab} | O | \psi'_{ab} \rangle$ and $\langle \psi_{ab} | O' | \psi_{ab} \rangle$. And, in general, there is no such relation:

$$\langle \psi'_{ab} | O | \psi'_{ab} \rangle = \langle \psi_{ab} | T_A^\dagger O \, T_A | \psi_{ab} \rangle \tag{9.62}$$

$$\langle \psi_{ab} | O' | \psi_{ab} \rangle = \langle \psi_{ab} | T_A O \, T_A^\dagger | \psi_{ab} \rangle \tag{9.63}$$

Again, only if $T_A^2 = I$, implying that T_A is both unitary and Hermitian, are the two expectation values equal. In our mirror transformation case, T_A satisfies this condition and we find

$$\langle \psi'_{ab} | X | \psi'_{ab} \rangle = \langle \psi_{ab} | X' | \psi_{ab} \rangle = \langle \psi_{ab} | Y | \psi_{ab} \rangle = \frac{L_2}{2} \tag{9.64}$$

$$\langle \psi'_{ab} | Y | \psi'_{ab} \rangle = \langle \psi_{ab} | Y' | \psi_{ab} \rangle = \langle \psi_{ab} | X | \psi_{ab} \rangle = \frac{L_1}{2} \tag{9.65}$$

$$\langle \psi'_{ab} | P_x | \psi'_{ab} \rangle = \langle \psi_{ab} | P_x' | \psi_{ab} \rangle = \langle \psi_{ab} | P_y | \psi_{ab} \rangle = 0 \tag{9.66}$$

$$\langle \psi'_{ab} | P_y | \psi'_{ab} \rangle = \langle \psi_{ab} | P_y' | \psi_{ab} \rangle = \langle \psi_{ab} | P_x | \psi_{ab} \rangle = 0 \tag{9.67}$$

where the first step in each line is the special relationship we have for the mirror transformation and the second step uses the relationships between the primed and unprimed operators we found for the passive case, again specific to the mirror transformation.

As we explained above, for the active case, we are interested in the transformed Hamiltonian, and we showed for the passive case what the transformed Hamiltonian is. We also have shown that the eigenstates of the transformed Hamiltonian are the transformed eigenstates of the untransformed Hamiltonian, which for this example implies that $|\psi'_{ab}\rangle$ is an eigenstate of H' with energy

$$E_{ab} = \frac{\hbar^2\,\pi^2}{2\,m}\left(\frac{a^2}{L_1^2} + \frac{b^2}{L_2^2}\right) \tag{9.68}$$

As we explained earlier, in general it is not true that $|\psi_E\rangle$ is an eigenstate of H' and also tht $|\psi'_E\rangle$ is an eigenstate of H. We see that same fact here: for $L_1 \neq L_2$, any eigenstate of H, with the box having L_1 along x and L_2 along y, simply cannot be an eigenstate of H' because L_1 runs along $x' = y$ and L_2 along $y' = x$ for H'.

Passive vs. Active Transformations

How are these two types of transformations related? Well, they are just different. Let us summarize what they do here:

▶ For a passive transformation, the states, the basis elements, and the operators are unchanged, but the transformation produces new basis elements and new operators. The transformed operators act in the transformed basis in the same way as the untransformed operators did in the untransformed basis. We work with the untransformed Hamiltonian, rewritten in terms of the transformed position and momentum operators $\{Q'\}$ and $\{P_q'\}$. We leave the states unchanged but decompose them in terms of the transformed basis.

▶ For an active transformation, we transform the states, the basis elements, and the operators. As for passive transformations, the transformed operators act in the transformed basis in the same way as the untransformed operators did in the untransformed basis. We work with the transformed Hamiltonian written in terms of the transformed operators. The decomposition of the transformed states in the transformed basis is the same as the decomposition of the untransformed states and operators in the untransformed basis.

Though they are different, the two transformations are related at a calculational level. Consider an active transformation T_A that transforms $\{|q\rangle\}$ into $\{|q'\rangle\}$ and the state $|\psi\rangle$ into $|\psi'\rangle$. Also, consider a passive transformation T_P that transforms the same $\{|q'\rangle\}$ into the same $\{|q\rangle\}$ and consider the same state $|\psi'\rangle$ under this transformation. The decomposition of the actively transformed state $|\psi'\rangle = T_A|\psi\rangle$ in terms of the untransformed basis $\{|q\rangle\}$ will be the same as the decomposition of the passively untransformed state $|\psi'\rangle$ in terms of the passively transformed basis $\{|q\rangle = T_P |q'\rangle\}$.

Symmetry Transformations

Definition

A coordinate transformation T is a symmetry transformation of a Hamiltonian H if it satisfies any of the three following equivalent statements

$$H' = T H T^{\dagger} = H \qquad \Longleftrightarrow \qquad T H = H T \qquad \Longleftrightarrow \qquad [T, H] = 0 \qquad (9.69)$$

That is, a symmetry transformation is a coordinate transformation that commutes with the Hamiltonian.

As we showed earlier (Equation 9.43, the commutativity of H and T ensures that the eigenstates $|\psi_E\rangle$ of H will now be eigenstates of H'. For active transformations, the eigenstates $|\psi_E'\rangle$ of H' will also be eigenstates of H. That proof also implied that the eigenvalue of a given eigenstate would be the same for H and H'.

More generally, we can conclude

$$\langle \psi | H' | \psi \rangle = \langle \psi | H | \psi \rangle \tag{9.70}$$

which is valid for passive and active transformations. For the latter, we also have

$$\langle \psi' | H | \psi' \rangle = \langle \psi' | H' | \psi' \rangle = \langle \psi | H | \psi \rangle \tag{9.71}$$

where the first step uses $H' = H$ and the second uses unitarity of T. Note the distinction between the full relation given above and simple unitarity.

Example 9.3: 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, $L_1 = L_2 \equiv L$. The symmetry of the potential ensures that $V(x, y) = V(y, x)$. The asymmetry of the potential under exchange of x and y was the reason that $H' \neq H$ before. Now, with the symmetric potential, we obtain

$$H' = T \left(\frac{P_x^2 + P_y^2}{2\,m} + V(X, Y) \right) T^\dagger = \frac{(P_x')^2 + (P_y')^2}{2\,m} + V(X', Y') \tag{9.72}$$

$$= \frac{P_y^2 + P_x^2}{2\,m} + V(Y', X') = \frac{P_x^2 + P_y^2}{2\,m} + V(X, Y) = H \tag{9.73}$$

where in the third step 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 eigenstates now have wavefunctions

$$\psi_{ab,q}(x, y) = \frac{2}{L} \sin\left(\frac{a\pi x}{L}\right) \sin\left(\frac{b\pi y}{L}\right) \tag{9.74}$$

$$\psi_{ab,q'}(x', y') = \frac{2}{L} \sin\left(\frac{b\pi x'}{L}\right) \sin\left(\frac{a\pi y'}{L}\right) \tag{9.75}$$

$$\psi'_{ab,q'}(x', y') = \frac{2}{L} \sin\left(\frac{a\pi x'}{L}\right) \sin\left(\frac{b\pi y'}{L}\right) \tag{9.76}$$

$$\psi'_{ab,q}(x, y) = \frac{2}{L} \sin\left(\frac{b\pi x}{L}\right) \sin\left(\frac{a\pi y}{L}\right) \tag{9.77}$$

We see that $\psi'_{ab,q'}(x', y') = \psi_{ba,q'}(x', y')$ and $\psi_{ab,q}(x, y) = \psi'_{ba,q}(x, y)$. The reversal of the indices is just a labeling issue; it does hold that $\psi_{ab,q'}(x', y')$ and $\psi'_{ab,q'}(x', y')$ have the same quantum number a in the x' direction and b in the y' direction, so the two states are the same. So the eigenstates of H and H' are the same and have the same energies, as is required by $H = H'$.

There is not much else to be said – all our other results hold, with $L_1 = L$ and $L_2 = L$. Because the condition for a coordinate transformation to be a symmetry transformation only involves the Hamiltonian, it can only modify results for H and the eigenstates.

Lecture 25:
Discrete and Continuous Coordinate and Symmetry
Transformations

Revision Date: 2007/12/05

Discrete and Continous Coordinate and Symmetry Transformations

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.

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.

Discrete and Continous Coordinate and Symmetry Transformations (cont.)

Example 9.4: Discrete Transformation Example: Parity Transformation

We have already gone through in some detail the mirror transformation example; it is a good example of a discrete transformation. But let's do the parity transformation also.

We define the parity transformation to be the operator that maps $|q\rangle$ to $|-q\rangle$; explicitly,

$$|-q\rangle = \Pi |q\rangle \tag{9.78}$$

The sign should be flipped on all the spatial dimensions. Note that this fully defines the transformation, as any state can be decomposed in terms of the $\{|q\rangle\}$ to determine how the transformation affects it, and any operator can be transformed according to the standard rule that we proved earlier,

$$O' = \Pi O \Pi^\dagger \tag{9.79}$$

Discrete and Continous Coordinate and Symmetry Transformations (cont.)

Let's consider the effect of an *active* parity transformation on a state. As we discussed above, the state changes. To compare the transformed state $|\psi'\rangle$ and untransformed state $|\psi\rangle$, we need to look at their representations in the same basis; let's pick the $\{|q\rangle\}$ basis. So:

$$\psi'_q(q) = \langle q|\psi'\rangle = \langle q|\Pi|\psi\rangle = \langle q|\Pi\sum_{q_1}|q_1\rangle\langle q_1|\psi\rangle \tag{9.80}$$

$$= \langle q|\Pi\sum_{q_1}\psi_q(q_1)|q_1\rangle = \langle q|\sum_{q_1}\psi_q(q_1)|-q_1\rangle = \sum_{q_1}\psi_q(q_1)\,\delta_{q,-q_1} \tag{9.81}$$

$$= \psi_q(-q) \tag{9.82}$$

We see that the effect of the transformation is to invert the wavefunction through the origin; or, thought of in terms of representation in terms of a basis, to exchange the q and $-q$ elements of the representation.

Since the parity transformation maps the position basis elements in the above way, and we can decompose the momentum basis elements in terms of position basis elements,

$$|p\rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} dq \, e^{\frac{i}{\hbar} p q} |q\rangle \tag{9.83}$$

we see that the action of the parity operator is

$$\Pi \, |p\rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} dq \, e^{\frac{i}{\hbar} p q} \, \Pi \, |q\rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} dq \, e^{\frac{i}{\hbar} p q} |-q\rangle \tag{9.84}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} dq_1 \, e^{-\frac{i}{\hbar} p q_1} |q_1\rangle = |-p\rangle \tag{9.85}$$

where in the third step we changed variables in the integration from q to $q_1 = -q$. This result makes sense: if you invert the coordinate system, the momenta should also get inverted.

What are the parity transformed operators? Formally, we have

$$O' = \Pi O \Pi^\dagger \tag{9.86}$$

This is not particularly useful until we use it to calculate matrix elements in the untransformed basis: We want to know what the parity-transformed operator is in terms of the untransformed basis elements; by definition, the parity transformed operator acts on the transformed basis elements in the same way that the untransformed operator acts on the untransformed basis elements (Equations 9.5 and 9.39). Let's calculate:

$$\langle q_1 | O' | q_2 \rangle = \langle q_1 | \Pi O \Pi^\dagger | q_2 \rangle = \langle -q_1 | O | -q_2 \rangle \tag{9.87}$$

How to interpret this now depends on the operator: how is $\langle -q_1 | O | -q_2 \rangle$ related to $\langle q_1 | O | q_2 \rangle$? For $O = Q$, the answer is simple

$$\langle -q_1 | Q | -q_2 \rangle = -q_1 \delta_{-q_1, -q_2} = -\left(q_1 \delta_{q_1, q_2} \right) = -\langle q_1 | Q | q_2 \rangle \tag{9.88}$$

$$\implies Q' = -Q \tag{9.89}$$

The same proof holds for P_q, so $P_q' = -P_q$. These results are not suprising.

One fine point: we had $\Pi^\dagger$ acting on $|q_2\rangle$ and Π acting to the left on $\langle q_1|$. The fact that $\Pi^\dagger|q_2\rangle = |-q_2\rangle$ and $\langle q_1|\Pi = \langle -q_1|$; *i.e.*, that Π and $\Pi^\dagger$ have simple actions on states; is specific to the parity operator. In general, it is not always so easy to express O' in terms of O. In our mirror inversion example, we saw that $X' = Y$ and $Y' = X$, but X' could not be written in terms of X and neither could Y' be in terms of Y.

Other interesting features of the parity operator:

- ▶ The parity operator is not only unitary, it is Hermitian. This is because $\Pi^2 = I$, so $\Pi^{-1} = \Pi$. Unitarity implies $\Pi^{-1} = \Pi^{\dagger}$, so we have $\Pi^{\dagger} = \Pi$ and hence Π is Hermitian.

- ▶ Π's eigenvalues are ± 1. Since Π is unitary, its eigenvalues are complex numbers $\{\omega\}$ of unity modulus. (Recall, we proved this in Section 3.6.) Since Π is Hermitian, its eigenvalues are also real. The only allowed eigenvalues are ± 1. (See next item for why both eigenvalues are found in practice.)

- ▶ Its eigenvectors are states whose position-space representations are even or odd under inversion through the origin. This is an obvious consequence of the above: the parity transformation changes the position space wavefunction from $\psi_q(q)$ to $\psi_q(-q)$, so even functions, for which $\psi_q(-q) = \psi_q(q)$ are eigenvectors with eigenvalue $+1$ and odd functions, for which $\psi_q(-q) = -\psi_q(q)$ are eigenvectors with eigenvalue -1.

When is a parity transformation a symmetry transformation?

Recall that the definition of a symmetry transformation is that the Hamiltonian must be invariant under the transformation, $H' = T H T^{\dagger}$. From our exercise above in calculating the effect of the transformation on the $\{Q\}$ and $\{P_q\}$ operators, we see that the Hamiltonian can only be invariant if it contains only even functions of the $\{Q\}$ and $\{P_q\}$. This will not hold for an arbitrary potential, but does hold for the free particle, the particle in a symmetric box, and the simple harmonic oscillator.

Recall that an alternate requirement is that H and Π commute. If the operators commute, they have simultaneous eigenvectors. Therefore, if $[H, \Pi] = 0$, the eigenstates of H will be parity eigenvectors – they must be even or odd. This can be turned into the condition that Π be a symmetry transformation of H: if the eigenvectors of H are even or odd, then they are eigenvectors of Π too and so $[H, \Pi] = 0$ must be true. Thus, the fact that the eigenstates of the free particle, symmetric box, and SHO Hamiltonians are even or odd implies that Π is a symmetry transformation of H in these situations.

The special role of parity transformations in our model of fundamental interactions.

Most spatial coordinate transformations, such as translations and rotations, are symmetry transformations of the Hamiltonians of the fundamental forces – electromagnetism, the strong force, the weak force, and gravity. Parity is a symmetry transformation of all but the weak force. What it means is that, though physical laws pick out no particular position or orientation in space as special, they do pick a particular handedness as special. Shankar goes through the discovery of parity violation in Section 11.4; it is worth reading. At a practical level, the way that parity violation is implemented in the Standard Model of particle physics is that right-handed particles do not interact via the weak force, only left-handed particles do.

Later, we will study the time-reversal transformation (which does not qualify as a coordinate transformation because of the way time is treated differently than space in non-relativistic quantum mechanics). We shall see that this transformation is also not a symmetry of the weak interaction Hamiltonian at a measurable level, indicating that the fundamental interactions also distinguish directions of time flow.

Discrete and Continous Coordinate and Symmetry Transformations (cont.)

General Properties of Continuous Coordinate and Symmetry Transformations

Consider a coordinate transformation that can be parameterized by and differentiated with respect to a continuous variable ε, $T = T(\varepsilon)$. This might be the translation vector for a translation, the rotation angle for a rotation, 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, $T(\varepsilon)$ for $\varepsilon \to 0$ – may always be written in the form

$$T(\varepsilon) = I - \frac{i}{\hbar}\,\varepsilon\,G \tag{9.90}$$

where G is some operator to be determined. This property arises simply because $T(\varepsilon) \to I$ 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.

Discrete and Continous Coordinate and Symmetry Transformations (cont.)

What may we say about G? Let's investigate the consequences of the unitarity of $T(\varepsilon)$. To do this, we need to know what $T^{-1}(\varepsilon)$ is? It is $T^{-1}(\varepsilon) = I + \frac{i}{\hbar} \varepsilon G$, which we can prove by applying it to $T(\varepsilon)$:

$$T^{-1}(\varepsilon) \, T(\varepsilon) = \left(I + \frac{i}{\hbar} \varepsilon G \right) \left(I - \frac{i}{\hbar} \varepsilon G \right) = I + \mathcal{O}(\varepsilon^2) = I \qquad (9.91)$$

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:

$$T^{\dagger}(\varepsilon) = T^{-1}(\varepsilon) \qquad (9.92)$$

$$I^{\dagger} + \frac{i}{\hbar} \, G^{\dagger} = I + \frac{i}{\hbar} \varepsilon \, G \qquad (9.93)$$

$$G^{\dagger} = G \qquad (9.94)$$

With some knowledge of the properties of G, we may construct the full transformation for arbitrary ε by taking an infinite product of infinitesimal transformations:

$$T(\varepsilon) = \lim_{N \to \infty} \left[T\left(\frac{\varepsilon}{N}\right) \right]^N = \lim_{N \to \infty} \left[I - \frac{i}{\hbar}\frac{\varepsilon}{N}\,G \right]^N = \exp\left(-\frac{i}{\hbar}\,\varepsilon\,G \right) \qquad (9.95)$$

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:

$$T^{-1}(\varepsilon) \left[i\,\hbar\,\frac{d}{d\varepsilon}\,T(\varepsilon) \right] = T^{-1}(\varepsilon)\,T(\varepsilon)\,G = G \qquad (9.96)$$

That is, G can be obtained by differentiating $T(\varepsilon)$. Due to the relation between G and $T(\varepsilon)$, G is called the generator of the coordinate transformation $T(\varepsilon)$. Because G is Hermitian, G is allowed to be an observable, a property that will be interesting for symmetry transformations.

We note that the explicit form for $T(\varepsilon)$ lets us write explicitly how position-basis elements, states, and operators transform:

$$|q'\rangle = \exp\left(-\frac{i}{\hbar}\,\varepsilon\,G\right)|q\rangle \qquad (9.97)$$

$$|\psi'\rangle = \exp\left(-\frac{i}{\hbar}\,\varepsilon\,G\right)|\psi\rangle \qquad (9.98)$$

$$O' = \exp\left(-\frac{i}{\hbar}\,\varepsilon\,G\right)O\exp\left(\frac{i}{\hbar}\,\varepsilon\,G\right) \qquad (9.99)$$

The effect of the transformation $T(\varepsilon)$ 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:

$$|g'\rangle = e^{-\frac{i}{\hbar}\,\varepsilon\,G}|g\rangle = e^{-\frac{i}{\hbar}\,\varepsilon\,g}|g\rangle \qquad (9.100)$$

That is, the eigenstates of G transform very simply under $T(\varepsilon)$: they retain the same direction in Hilbert space, picking up only a unity-modulus factor.

The above suggests that an arbitrary state's evolution may be calculated simply if we work with the $\{|g\rangle\}$-basis representation of the state. Denote by $\psi_g(g) = \langle g | \psi \rangle$ the $\{|g\rangle\}$-basis representation ("$\{|g\rangle\}$-basis wavefunction") of $|\psi\rangle$. Then we have

$$\psi_g'(g) = \langle g | \psi' \rangle = \langle g | e^{-\frac{i}{\hbar}\,\varepsilon\,G} | \psi \rangle = e^{-\frac{i}{\hbar}\,\varepsilon\,g}\,\langle g | \psi \rangle = e^{-\frac{i}{\hbar}\,\varepsilon\,g}\,\psi_g(g) \qquad (9.101)$$

That is, 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:

$$\langle g_1 | O' | g_2 \rangle = \langle g_1 | e^{-\frac{i}{\hbar}\,\varepsilon\,G} O\, e^{\frac{i}{\hbar}\,\varepsilon\,G} | g_2 \rangle = e^{-\frac{i}{\hbar}\,\varepsilon\,(g_1 - g_2)}\,\langle g_1 | O | g_2 \rangle \qquad (9.102)$$

Again, we get a very simple relation between the matrix elements of O' and those of O in the $\{|g\rangle\}$ basis.

Discrete and Continous Coordinate and 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 is an important statement: it implies that G and H are simultaneously diagonalizable. This has a number of important implications:

▶ The eigenvectors of H are eigenvectors of G and vice versa.

▶ G commutes with the unitary time evolution operator $U = e^{\frac{i}{\hbar} H t}$, $[G, U(t)] = 0$.

▶ 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:

$$\langle \psi(t) | G | \psi(t) \rangle = \langle \psi(0) | U^\dagger(t) \, G U(t) | \psi(0) \rangle \tag{9.103}$$

$$= \langle \psi(0) | U^\dagger(t) \, U(t) \, G | \psi(0) \rangle = \langle \psi(0) | G | \psi(0) \rangle \tag{9.104}$$

That is, in any way one can think of, the physical quantity associated with G is a conserved quantity.

Example 9.5: Spatial Translation Transformation of the Free Particle

The archetypal continuous coordinate transformation is spatial translation. We will consider spatial translations acting on a one-dimensional free particle system. The Hamiltonian is, obviously,

$$H = \frac{P^2}{2\,m} \tag{9.105}$$

We will consider the specific transformation

$$x' = x - a \tag{9.106}$$

The transformation acts on the position basis kets as

$$T(a)\,|x = u\rangle = |x' = u\rangle = |x = u + a\rangle \tag{9.107}$$

That is, the transformation operator translates the position basis kets to the right by a.

Discrete and Continous Coordinate and Symmetry Transformations (cont.)

Let's consider some arbitrary state $|\psi\rangle$, not necessarily an eigenstate of H. First, let's write the position-basis representation of the state in terms of the untransformed position basis:

$$\langle x = u \,|\psi\rangle = \psi_x(u) \tag{9.108}$$

Let's now write the representation in terms of the transformed basis:

$$\langle x' = u \,|\psi\rangle = \langle x = u + a \,|\psi\rangle = \psi_x(u + a) \tag{9.109}$$

$$\text{or} \qquad \psi_{x'}(x' = u) = \psi_x(x = u + a) \tag{9.110}$$

That is, the untransformed state at $x' = u$ is given by the untransformed state at $x = u + a$, not surprisingly.

What are the position-basis representations of the actively transformed state, $|\psi'\rangle$? Recall that unitarity tells us

$$\langle x' = u \,|\, \psi'\,\rangle = \langle x = u \,|\, \psi \,\rangle \qquad \Longrightarrow \qquad \psi'_{x'}(x' = u) = \psi_x(x = u) \qquad (9.111)$$

That is, the transformed state looks, relative to the transformed coordinate system, as the untransformed state does relative to the untransformed coordinate system. The transformed state in the untransformed coordinate system is again easily obtained:

$$\langle x' = u \,|\, \psi'\,\rangle = \langle x = u + a \,|\, \psi'\,\rangle = \psi'_x(x = u + a) \qquad (9.112)$$

where the first step comes from the transformation rule and the second from the definition of the x-basis representation. Hence

$$\psi'_x(x = u + a) = \psi'_{x'}(x' = u) = \psi_x(x = u) \qquad (9.113)$$

So, again not surprisingly, when both are viewed relative to the untransformed coordinate system, the transformed state at $u + a$ is the same as the untransformed state at u – translation of the state by a.

Now, let's look at operators. As usual, we use the relation between operator matrix elements to sort this out. That is, from $X' = T X T^\dagger$ we have

$$\langle x' = u |X'|x' = v \rangle = \langle x = u |X |x = v \rangle = u\,\delta(u - v) \tag{9.114}$$

Similarly

$$\langle x' = u |P'|x' = v \rangle = \langle x = u |P |x = v \rangle = -i\,\hbar\,\frac{d}{du}\,\delta(u - v) \tag{9.115}$$

We are usually more interested in writing the transformed operators in terms of the untransformed operators, so let's do that:

$$\langle x' = u | X' | x' = v \rangle = u \, \delta(u - v) = \langle x = u + a | (X - I\,a) | x = v + a \rangle \qquad (9.116)$$

where we have used $|x' = u\rangle = |x = u + a\rangle$ and similarly for $|x' = v\rangle$. We had to subtract off the $I\,a$ term so that the right side would equal the center expression; otherwise, we would have obtained $(u + a)\,\delta(u - v)$. The kets and bras being used on the two sides are equal and arbitrary, so we have

$$X' = X - I\,a \qquad (9.117)$$

This is as expected: if you calculate the x' expectation value of a state, it is a less than the x expectation value because of the translation.

P comes out as follows:

$$\langle x' = u | P' | x' = v \rangle = -i\hbar \frac{d}{du} \delta(u-v) = \langle x = u+a | P | x = v+a \rangle \quad (9.118)$$

because the P matrix element does not care about the actual value of u or v, only the difference between the two. So, we have

$$P' = P \quad (9.119)$$

Finally, let us look at H. From the above results, we immediately have

$$H' = \frac{(P')^2}{2\,m} = \frac{P^2}{2\,m} = H \quad (9.120)$$

So, translation is a symmetry transformation of H.

Discrete and Continous Coordinate and Symmetry Transformations (cont.)

With our discussion of continuous transformations, we are now inclined to ask: what is the generator of the translation transformation? The difficulty is that we have no explicit form for $T(a)$; we need one. This we can most easily get by considering the infinitesimal version of $T(a)$. We have that

$$T(\delta a) = I - \frac{i}{\hbar} \delta a \, G \qquad (9.121)$$

So,

$$\langle x = u \,|\, T(\delta a) \,|\, \psi \rangle = \langle x = u \,|\, \psi \rangle - \delta a \, \frac{i}{\hbar} \langle x = u \,|\, G \,|\, \psi \rangle \qquad (9.122)$$

$$= \psi_x(x = u) - \delta a \, \frac{i}{\hbar} \langle x = u \,|\, G \,|\, \psi \rangle \qquad (9.123)$$

Discrete and Continous Coordinate and Symmetry Transformations (cont.)

We also have that

$$\langle x = u \,|\, T(\delta a) \,|\psi\rangle = \langle x = u - a \,|\psi\rangle = \psi_x(x = u - a) \qquad (9.124)$$

$$= \psi_x(x = u) - \delta a \,\frac{d}{dx}\, \psi_x(x) \qquad (9.125)$$

where in the first step we acted to the left with $T(\delta a)$. So,

$$-\frac{i}{\hbar} \,\langle x = u \,|G|\psi\rangle = -\frac{d}{dx}\, \psi_x(x) \quad \Longleftrightarrow \quad \langle x = u \,|G|\psi\rangle = -i\,\hbar\,\frac{d}{dx}\, \psi_x(x) \quad (9.126)$$

We know that

$$\langle x = u \,|P|\psi\rangle = -i\,\hbar\,\frac{d}{dx}\, \psi_x(x) \qquad (9.127)$$

and, since u and $|\psi\rangle$ are arbitrary, it therefore holds

$$G = P \qquad \Longleftrightarrow \qquad T(\delta a) = I - \frac{i}{\hbar}\,\delta a\,P \qquad \Longleftrightarrow \qquad T(a) = e^{-\frac{i}{\hbar}\,a\,P} \qquad (9.128)$$

When $T(a)$ is a symmetry transformation of H, then the corresponding generator commutes with H, so we have just performed a very complicated proof of $[H, P] = 0$ for $H = P^2/2\,m$. However, we now see that the fact that these two operators commute is more fundamental than just the fact that H is a function of P alone; we see that it is more deeply a result of the fact that H is invariant under translations.

We note that, now that we have an explicit form for $T(a)$, we can explicitly calculate the transformation properties for position-basis elements, states, and operators:

$$|x'\rangle = \exp\left(-\frac{i}{\hbar} a P\right)|x\rangle = \int_{-\infty}^{\infty} dp \, \exp\left(-\frac{i}{\hbar} a P\right)|p\rangle\langle p|x\rangle \tag{9.129}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dp \, \exp\left(-\frac{i}{\hbar} a p\right)|p\rangle e^{-\frac{i}{\hbar} p x} \tag{9.130}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dp \, |p\rangle e^{-\frac{i}{\hbar} p(x+a)} = |x+a\rangle \tag{9.131}$$

$$|\psi'\rangle = \exp\left(-\frac{i}{\hbar} a P\right)|\psi\rangle = \int_{-\infty}^{\infty} dp \, \exp\left(-\frac{i}{\hbar} a P\right)|p\rangle\langle p|\psi\rangle \tag{9.132}$$

$$= \int_{-\infty}^{\infty} dp \, \exp\left(-\frac{i}{\hbar} a p\right)|p\rangle\psi_p(p) = \int_{-\infty}^{\infty} dp \, e^{-\frac{i}{\hbar} a p}\psi_p(p)|p\rangle \tag{9.133}$$

To see that this is shifted, we need to project onto $|x\rangle$:

$$\langle x|\psi'\rangle = \int_{-\infty}^{\infty} dp\, e^{-\frac{i}{\hbar}ap}\psi_p(p)\langle x|p\rangle = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dp\, e^{-\frac{i}{\hbar}ap}\psi_p(p)\, e^{\frac{i}{\hbar}px} \quad (9.134)$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dp\,\psi_p(p)\, e^{\frac{i}{\hbar}p(x-a)} = \int_{-\infty}^{\infty} dp\,\langle x-a|p\rangle\langle p|\psi\rangle = \langle x-a|\psi\rangle \quad (9.135)$$

That is, $\psi_x'(x) = \psi_x(x-a)$ as expected. The operators obey

$$X' = e^{-\frac{i}{\hbar}aP} X\, e^{\frac{i}{\hbar}aP} \quad (9.136)$$

$$= e^{-\frac{i}{\hbar}aP}\left[e^{\frac{i}{\hbar}aP}X + e^{\frac{i}{\hbar}aP}\left[X,\frac{i}{\hbar}aP\right]\right] \quad (9.137)$$

$$= X - a\,I \quad (9.138)$$

as expected. We used $[A, e^B] = e^B [A, B]$ (you can check that it is true by direct calculation). It is trivial that $P' = P$ and $H' = H$ because both consist of only the operator P, which we know commutes with $T(a)$.

Finally, we consider the effect of spatial translations when we write everything in terms of the momentum basis, which is the eigenbasis of the generator P of translations:

$$|p'\rangle = e^{-\frac{i}{\hbar} a P}|p\rangle = e^{-\frac{i}{\hbar} a p}|p\rangle \tag{9.139}$$

$$\psi_p' = \langle p|\psi'\rangle = \langle p|e^{-\frac{i}{\hbar} a P}|\psi\rangle = e^{-\frac{i}{\hbar} a p}\langle p|\psi\rangle = e^{-\frac{i}{\hbar} a p}\psi_p(p) \tag{9.140}$$

$$\langle p_1|O'|p_2\rangle = \langle p_1|e^{-\frac{i}{\hbar} a P}O e^{\frac{i}{\hbar} a P}|p_2\rangle = e^{-\frac{i}{\hbar} a(p_1-p_2)}\langle p_1|O|p_2\rangle \tag{9.141}$$

As we saw for the general case, things simplify substantially in this basis.

Lecture 26:
Time Transformations
Classical Limit

Revision Date: 2007/12/07

Time Transformations

Time Translation vs. Coordinate Transformations

At a fundamental level, the time coordinate is different from the spatial coordinates in non-relativistic quantum mechanics. The main reason is that there is no observable associated with time. There is no momentum conjugate to time with which to construct a 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

$$T(t) \, |q(t=0) \, \rangle = |q' \, \rangle = |q(t) \, \rangle \tag{9.142}$$

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.

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,

$$T(t) = U(t) = \mathcal{T}\left[\exp\left(-\frac{i}{\hbar}\int_0^t dt' \, H(t')\right)\right] = \lim_{N\to\infty} \prod_{j=0}^{N-1} \exp\left[-\frac{i}{\hbar}\left(\frac{t}{N}\right)H\left(\frac{j}{N}t\right)\right]$$

(9.143)

(Equation 4.35) 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

$$|\psi(t)\rangle = |\psi'\rangle = U(t)\,|\psi(t=0)\rangle$$

(9.144)

For an arbitrary operator, we take as a requirement

$$\langle \psi_1(t) \,|\, O(t) \,|\, \psi_2(t) \rangle \equiv \langle \psi_1' \,|\, O' \,|\, \psi_2' \rangle = \langle \psi_1(t=0) \,|\, O_0 \,|\, \psi_2(t=0) \rangle \tag{9.145}$$

$$\Longleftrightarrow \qquad O(t) \equiv O' = U(t) \, O_0 \, U^\dagger(t) \tag{9.146}$$

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:

$$\langle \psi_1(t) \,|\, O_0 \,|\, \psi_2(t) \rangle = \langle \psi_1(t=0) \,|\, U^\dagger(t) \, O_0 \, U(t) \,|\, \psi_2(t=0) \rangle \tag{9.147}$$

$$\langle \psi_1(t=0) \,|\, O(t) \,|\, \psi_2(t=0) \rangle = \langle \psi_1(t=0) \,|\, U(t) \, O_0 \, U^\dagger(t) \,|\, \psi_2(t=0) \rangle \tag{9.148}$$

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

$$U(t + dt, t) = 1 - \frac{i}{\hbar} H(t)\, dt \tag{9.149}$$

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,

$$U(t) = e^{-\frac{i}{\hbar} H t} \tag{9.150}$$

We see that our coordinate transformation formulae carry through for time translation even though the analogy between time translations and coordinate transformations is not perfect:

$$|q'\rangle \equiv |q(t)\rangle = e^{-\frac{i}{\hbar} H t}|q(t=0)\rangle \equiv e^{-\frac{i}{\hbar} H t}|q\rangle \tag{9.151}$$

$$|\psi'\rangle \equiv |\psi(t)\rangle = e^{-\frac{i}{\hbar} H t}|\psi(t=0)\rangle \equiv e^{-\frac{i}{\hbar} H t}|\psi\rangle \tag{9.152}$$

$$O(t) \equiv O' = e^{-\frac{i}{\hbar} H t}\, O\, e^{\frac{i}{\hbar} H t} \tag{9.153}$$

As we did for coordinate transformations, we are led to consider the effect of the transformation on the eigenstates of the generator:

$$|E(t)\rangle \equiv |E'\rangle = e^{-\frac{i}{\hbar}Ht}|E\rangle = e^{-\frac{i}{\hbar}Et}|E\rangle \qquad (9.154)$$

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

$$\langle E|\psi(t)\rangle \equiv \langle E|\psi'\rangle = \langle E|e^{-\frac{i}{\hbar}Ht}|\psi\rangle = e^{-\frac{i}{\hbar}Et}\langle E|\psi\rangle = e^{-\frac{i}{\hbar}Et}\psi_E(E) \quad (9.155)$$

Consider the time transformation of matrix elements of the operators in the energy eigenbasis:

$$\langle E_1 | O(t) | E_2 \rangle = \langle E_1 | O' | E_2 \rangle = \langle E_1 | e^{-\frac{i}{\hbar} H t} O e^{\frac{i}{\hbar} H t} | E_2 \rangle = e^{-\frac{i}{\hbar}(E_1 - E_2) t} \langle E_1 | O | E_2 \rangle \tag{9.156}$$

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

$$\Pi_t|\psi(t)\rangle = |\psi(t)\rangle \tag{9.157}$$

Let's test this by asking what properties the Hamiltonian would have to have in order for time reversal to be a symmetry transformation of H. 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:

$$i\hbar \frac{d}{dt}|\phi(t)\rangle = i\hbar \frac{d}{dt}|\psi(-t)\rangle = -i\hbar \frac{d}{d(-t)}|\psi(-t)\rangle = -i\hbar \frac{d}{dt'}|\psi(t')\rangle \tag{9.158}$$

$$= -H(t')|\psi(t')\rangle = -H(-t)|\psi(-t)\rangle = -H(-t)|\phi(t)\rangle \tag{9.159}$$

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

$$i\,\hbar\,\frac{d}{dt}\,\psi_x(x,t) = H\left(x, -i\,\hbar\,\frac{d}{dx}, t\right)\,\psi_x(x,t) \tag{9.160}$$

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:

$$-i\,\hbar\,\frac{d}{dt}\,\psi_x^*(x,t) = H^*\left(x, -i\,\hbar\,\frac{d}{dx}, t\right)\,\psi_x^*(x,t) \tag{9.161}$$

Now, define $\phi_x(x, t) = \psi_x^*(x, -t)$. Then we have

$$-i\,\hbar\,\frac{d}{dt}\,\phi_x(x, -t) = H^*\left(x, -i\,\hbar\,\frac{d}{dx}, t\right)\,\phi_x(x, -t) \tag{9.162}$$

Change variables to $t' = -t$:

$$-i\,\hbar\,\frac{d}{d(-t')}\,\phi_x(x, t') = H^*\left(x, -i\,\hbar\,\frac{d}{dx}, -t'\right)\,\phi_x(x, t') \tag{9.163}$$

$$i\,\hbar\,\frac{d}{dt'}\,\phi_x(x, t') = H^*\left(x, -i\,\hbar\,\frac{d}{dx}, -t'\right)\,\phi_x(x, t') \tag{9.164}$$

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

$$\Pi_t |\psi(t)\,\rangle = |\phi(t)\,\rangle \quad \text{with} \quad \phi_x(x,t) = \psi_x^*(x,-t) \Longleftrightarrow \langle x\,|\phi(t)\,\rangle = \langle \psi(-t)\,|x\,\rangle \quad (9.165)$$

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.

Section 10
Classical Limit

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 $\hbar \to 0$. We can see this explicitly:

$$\frac{d}{dt}\langle\Omega\rangle = \left(\frac{d}{dt}\langle\psi\,|\right)\Omega\,|\psi\rangle + \langle\psi\,|\Omega\left(\frac{d}{dt}\,|\psi\rangle\right) + \langle\psi\,|\left(\frac{d\Omega}{dt}\right)|\psi\rangle \tag{10.1}$$

$$= -\frac{i}{\hbar}\left[-\left(\langle\psi\,|H\rangle\,\Omega\,|\psi\rangle + \langle\psi\,|\Omega\,(H|\psi\rangle)\right)\right] + \langle\psi\,|\left(\frac{d\Omega}{dt}\right)|\psi\rangle \tag{10.2}$$

$$= -\frac{i}{\hbar}\langle\psi\,|[\Omega, H]|\psi\rangle + \langle\psi\,|\left(\frac{d\Omega}{dt}\right)|\psi\rangle \tag{10.3}$$

$$= -\frac{i}{\hbar}\langle[\Omega, H]\rangle + \left\langle\frac{d\Omega}{dt}\right\rangle \tag{10.4}$$

If the observable Ω has no explicit time-dependence, this reduces to

$$\frac{d}{dt} \langle \Omega \rangle = -\frac{i}{\hbar} \langle [\Omega, H] \rangle \tag{10.5}$$

This result is known as the Ehrenfest Theorem.

Applications of the Ehrenfest Theorem

To make use of it, let's consider some examples. For $\Omega = P$, we have

$$\frac{d}{dt}\langle P \rangle = -\frac{i}{\hbar}\langle [P, H] \rangle \tag{10.6}$$

We know P commutes with $P^2/2m$, so the only interesting term will be $[P, V(X)]$. We can see what that is via Taylor expansion of $V(X)$:

$$[P, V(X)] = \left[P, \sum_{n=0}^{\infty} V_n X^n \right] = \sum_{n=0}^{\infty} V_n [P, X^n] \tag{10.7}$$

$$= \sum_{n=0}^{\infty} V_n \left(-i\hbar n X^{n-1} \right) = -i\hbar \frac{dV(X)}{dX} \tag{10.8}$$

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

$$\frac{d}{dt}\langle P \rangle = -\left\langle \frac{dV}{dX} \right\rangle \tag{10.9}$$

which is the standard classical result for Newton's second law in the presence of a potential $V(x)$.

Another classical result is found by setting $\Omega = X$:

$$\frac{d}{dt}\langle X \rangle = -\frac{i}{\hbar} \left\langle [X, \frac{P^2}{2\,m}] \right\rangle \tag{10.10}$$

By the same kind of calculation as applied above, $[X, P^2] = 2\,i\,\hbar\,P$, so

$$\frac{d}{dt}\langle X \rangle = \frac{\langle P \rangle}{m} \tag{10.11}$$

which is the classical relation between velocity and linear momentum.

The above two results can be rewritten as

$$\frac{d}{dt}\langle P \rangle = -\left\langle \frac{dH}{dX} \right\rangle \qquad \frac{d}{dt}\langle X \rangle = \left\langle \frac{dH}{dP} \right\rangle \qquad (10.12)$$

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.

Lecture 27:
Analogies to Classical Mechanics

Revision Date: 2007/12/07

Correspondences between Hamiltonian Mechanics and Quantum Mechanics

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

$$\{f, g\}_{q,p} = \frac{\partial f}{\partial x} \frac{\partial g}{\partial p} - \frac{\partial g}{\partial x} \frac{\partial f}{\partial p} \tag{10.13}$$

We have in classical mechanics the generic result for the time evolution of a function F of the state variables x, p

$$\frac{df}{dt} = \{f, \mathcal{H}\} + \frac{\partial f}{\partial t} \tag{10.14}$$

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 $\frac{1}{i\hbar}[F, H]$, one obtains Ehrenfest's Theorem.

Correspondences between Hamiltonian Mechanics 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 $f = p$:

$$\frac{dx}{dt} = \{x, H\} = \frac{\partial H}{\partial p} \qquad \frac{dp}{dt} = \{p, H\} = -\frac{\partial H}{\partial x} \qquad (10.15)$$

We saw earlier the quantum version of these equations, again with $\{\ ,\ \}$ replaced by $\frac{1}{i\hbar}[\ ,\ ]$.

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

$$\frac{dp_q}{dt} = -\frac{\partial \mathcal{H}}{\partial q} \qquad \frac{dq}{dt} = \frac{\partial \mathcal{H}}{\partial p_q} \tag{10.16}$$

Consider the contact transformation defined by

$$q' = q'(q(t), p_q(t), t) \qquad p_q' = p_q'(q(t), p_q(t), t) \tag{10.17}$$

(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.)

Correspondences between Hamiltonian Mechanics and Quantum Mechanics (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,

$$\frac{dp'_q}{dt} = -\frac{\partial \mathcal{H}}{\partial q'} \qquad \frac{dq'}{dt} = \frac{\partial \mathcal{H}}{\partial p'_q} \tag{10.18}$$

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:

$$\{q', p'_q\}_{q,p_q} = \frac{\partial q'}{\partial q} \frac{\partial p'_q}{\partial p_q} - \frac{\partial p'_q}{\partial q} \frac{\partial q'}{\partial p_q} = 1 \tag{10.19}$$

The quantum analogue is Postulate 2, whose alternate version (see Section 6.5) is

$$[Q, P_q] = i\,\hbar \tag{10.20}$$

If we carry the Poisson bracket over to a commutator in the standard fashion, the classical requirement becomes

$$[Q', P_q'] = i\,\hbar \tag{10.21}$$

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:

$$[Q', P'] = Q'P' - P'Q' = T\,Q\,T^{\dagger}\,T\,P_q\,T^{\dagger} - T\,P_q\,T^{\dagger}\,T\,Q\,T^{\dagger} \tag{10.22}$$

$$= T\,Q\,P_q\,T^{\dagger} - T\,P_q Q\,T^{\dagger} = T\,[Q, P_q]\,T^{\dagger} = T\,i\,\hbar\,T^{\dagger} = i\,\hbar \tag{10.23}$$

Correspondences between Hamiltonian Mechanics and Quantum Mechanics (cont.)

Infinitesimal Classical and Quantum Transformations

An infinitesimal canonical transformation is of the form

$$q' = q + \delta\varepsilon \frac{\partial G}{\partial p_q} = q + \delta\varepsilon \{q, G\}_{q,p_q} \qquad p_q' = p_q - \delta\varepsilon \frac{\partial G}{\partial q} = p_q + \delta\varepsilon \{p_q, G\}_{q,p_q}$$

(10.24)

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:

$$\{q', p_q'\}_{q,p_q} = \frac{\partial q'}{\partial q} \frac{\partial p_q'}{\partial p_q} - \frac{\partial p_q'}{\partial q} \frac{\partial q'}{\partial p_q}$$

(10.25)

$$= \left(1 + \delta\varepsilon \frac{\partial^2 G}{\partial q \, \partial p_q}\right)\left(1 - \delta\varepsilon \frac{\partial^2 G}{\partial p_q \, \partial q}\right) - \left(-\delta\varepsilon \frac{\partial^2 G}{\partial q^2}\right)\left(\delta\varepsilon \frac{\partial^2 G}{\partial p^2}\right)$$

$$= 1 + \delta\varepsilon \left(\frac{\partial^2 G}{\partial q \, \partial p_q} - \frac{\partial^2 G}{\partial p_q \, \partial q}\right) + \delta\varepsilon^2 \left(-\frac{\partial^2 G}{\partial q \, \partial p_q} \frac{\partial^2 G}{\partial p_q \, \partial q} + \frac{\partial^2 G}{\partial q^2} \frac{\partial^2 G}{\partial p_q^2}\right)$$

With our condition on $G(q, p_q)$, the coefficient of $\delta\varepsilon$ vanishes. We neglect the term of order $(\delta\varepsilon)^2$, leaving

$$\{q', p_q'\}_{q, p_q} = 1 \tag{10.26}$$

The infinitesimal transformation may be rewritten as

$$\delta q = q' - q = \delta\varepsilon \, \{q, G\}_{q, p_q} \qquad \delta p_q = p_q' - p_q = \delta\varepsilon \, \{p, G\}_{q, p_q} \tag{10.27}$$

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

$$T(\delta\varepsilon) = I + \frac{i}{\hbar} \, \delta\varepsilon \, G \tag{10.28}$$

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 infinitesimal transformation of an arbitrary operator O is

$$O' = T(\delta\varepsilon)\, O\, T^{\dagger}(\delta\varepsilon) = \left(I + \frac{i}{\hbar}\, \delta\varepsilon\, G \right) O \left(I - \frac{i}{\hbar}\, \delta\varepsilon\, G \right) = O + \frac{\delta\varepsilon}{i\,\hbar}\, [O, G] + \mathcal{O}(\delta\varepsilon^2) \tag{10.29}$$

$$\text{or} \quad \delta O = O' - O = \frac{\delta\varepsilon}{i\,\hbar}\, [O, G] \tag{10.30}$$

$$\text{So,} \quad \delta Q = Q' - Q = \frac{\delta\varepsilon}{i\,\hbar}\, [Q, G] \qquad \delta P = P'_q - P_q = \frac{\delta\varepsilon}{i\,\hbar}\, [P_q, G] \tag{10.31}$$

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

$$\delta q = \delta\varepsilon \, \{q, p_q\}_{q,p_q} = \delta\varepsilon \qquad\qquad \delta p = \delta\varepsilon \, \{p_q, p_q\}_{q,p_q} = 0 \qquad (10.32)$$

$$\delta Q = \frac{\delta\varepsilon}{i\,\hbar} \, [Q, P_q] = \delta\varepsilon \qquad\qquad \delta P_q = \frac{\delta\varepsilon}{i\,\hbar} \, [P_q, P_q] = 0 \qquad (10.33)$$

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 $\delta\varepsilon$.

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, $\langle Q\rangle \to \langle Q\rangle + \delta\varepsilon$ is completely consistent with $Q' = Q - \delta\varepsilon\, I$ as well as with the classical transformation $q \to q + \delta\varepsilon$.

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.

Correspondences between Hamiltonian Mechanics and Quantum Mechanics (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

$$\delta q = \delta\varepsilon \left\{ q, \mathcal{H} \right\}_{q,p_q} \qquad\qquad \delta p = \delta\varepsilon \left\{ p_q, \mathcal{H} \right\}_{q,p_q} \qquad\qquad (10.34)$$

If we take $\delta\varepsilon = \delta t$, we have

$$\frac{\delta q}{\delta t} = \delta\varepsilon \left\{ q, \mathcal{H} \right\}_{q,p_q} \qquad\qquad \frac{\delta p}{\delta t} = \delta\varepsilon \left\{ p_q, \mathcal{H} \right\}_{q,p_q} \qquad\qquad (10.35)$$

which are just Hamilton's equations. For the quantum case, we have

$$\delta Q = \frac{\delta\varepsilon}{i\,\hbar} \left[Q, H \right] \qquad\qquad \delta P_q = \frac{\delta\varepsilon}{i\,\hbar} \left[P_q, H \right] \qquad\qquad (10.36)$$

Again, let $\delta\varepsilon = \delta t$ and also take expectation values. We thus have

$$\left\langle \frac{\delta Q}{\delta t} \right\rangle = \frac{1}{i\,\hbar} \left\langle [Q, H] \right\rangle \qquad\qquad \left\langle \frac{\delta P_q}{\delta t} \right\rangle = \frac{1}{i\,\hbar} \left\langle [P_q, H] \right\rangle \qquad\qquad (10.37)$$

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

$$O' = e^{\frac{i}{\hbar} \varepsilon G} \, O \, e^{-\frac{i}{\hbar} \varepsilon G} \tag{10.38}$$

We may rewrite this as follows:

$$O' = O + \frac{\varepsilon}{i\,\hbar}[O, G] + \frac{1}{2!} \left(\frac{\varepsilon}{i\,\hbar} \right)^2 [[O, G], G] + \cdots \tag{10.39}$$

(You can check that this is correct by simply expanding it.)

Correspondences between Hamiltonian Mechanics and Quantum Mechanics (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

$$\delta O = \delta \varepsilon \{O, G\}_{q,p_q} \qquad \Longleftrightarrow \qquad \frac{d}{d\varepsilon} O(q, p_q; \varepsilon) = \{O, G\}_{q,p_q} \qquad (10.40)$$

This holds for any function $O(q, p_q; \varepsilon)$, so apply it to $dO/d\varepsilon$:

$$\frac{d^2}{d\varepsilon} O(q, p_q; \varepsilon) = \frac{d}{d\varepsilon} \frac{dO}{d\varepsilon} = \left\{ \frac{dO}{d\varepsilon}, G \right\}_{q,p_q} = \left\{ \{O, G\}_{q,p_q}, G \right\}_{q,p_q} \qquad (10.41)$$

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:

$$O(q, p_q; \varepsilon) = O(q, p_q; \varepsilon = 0) + \varepsilon \{O, G\}_{q,p_q} + \frac{1}{2!} \varepsilon^2 \{\{O, G\}_{q,p_q}, G\}_{q,p_q} \qquad (10.42)$$

So, again, we have direct correspondence with the substitution $\{ \, , \, \} \rightarrow \frac{1}{i\hbar} [\, , \,]$.

Correspondences between Hamiltonian Mechanics and Quantum Mechanics (cont.)

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):

$$[H, G] = 0 \qquad \Longleftrightarrow \qquad H' = H \qquad (10.43)$$

We see from the previous page that the classical Hamiltonian $\mathcal{H}$ is unchanged if a similar condition holds involving Poisson brackets:

$$\{H, G\}_{q, p_q} = 0 \qquad \Longleftrightarrow \qquad \mathcal{H}' = \mathcal{H} \qquad (10.44)$$

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 11
Semiclassical Limit

Lecture 28:
Semiclassical (WKB) Approximation

Revision Date: 2008/01/10

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

$$\psi(x) = \psi(x = 0)\, e^{\pm \frac{i}{\hbar} p\, x} \qquad p = \sqrt{2\, m\, (E - V)} \tag{11.1}$$

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.:

$$\psi(x) = \begin{cases} \psi(x = 0)\, e^{\pm \frac{i}{\hbar} p_0 x} & 0 \le x < x_1 \\ \psi(x = x_1)\, e^{\pm \frac{i}{\hbar} p_1 (x - x_1)} & x_1 \le x < x_2 \\ \psi(x = x_2)\, e^{\pm \frac{i}{\hbar} p_2 (x - x_2)} & x_2 \le x < x_3 \\ \qquad \vdots & \qquad \vdots \end{cases} \tag{11.2}$$

We may write the above solution in the following less cumbersome form

$$\psi(x) = \psi(x = 0)\, \exp\left(\pm \frac{i}{\hbar} \int_0^x dx'\, p(x') \right) \qquad \text{where} \quad p(x) = \sqrt{2\, m\, (E - V(x))} \tag{11.3}$$

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.

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:

$$\left| \frac{1}{\lambda} \left(\frac{d\lambda}{dx} \lambda \right) \right| \ll 1 \quad \implies \quad \left| \frac{d\lambda}{dx} \right| \ll 1 \tag{11.4}$$

Derivation (for Unbound States)

Let's prove the above more rigorously. Our expected solution is of the form

$$\psi(x) = A\, e^{\frac{i}{\hbar}\phi(x)} \tag{11.5}$$

where $\phi(x)$ may be complex to avoid restricting the form of the solution. The Schrödinger Equation for our problem is

$$-\frac{\hbar^2}{2\,m}\frac{d^2}{dx^2}\,\psi(x) + V(x)\,\psi(x) = E\,\psi(x) \tag{11.6}$$

which we may rewrite as

$$\left[\frac{d^2}{dx^2} + \frac{p^2(x)}{\hbar^2}\right]\psi(x) = 0 \tag{11.7}$$

Now, inserting our expected solution, we have

$$\left(\frac{1}{\hbar}\frac{d\phi}{dx}\right)^2 - \frac{i}{\hbar}\frac{d^2\phi}{dx^2} - \frac{p^2}{\hbar^2} = 0 \tag{11.8}$$

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

$$2\pi = \frac{1}{\hbar}\left[\phi(x+\lambda) - \phi(x)\right] \tag{11.9}$$

$$2\pi\hbar \approx \lambda \frac{d\phi}{dx} \tag{11.10}$$

$$\lambda(x) \approx \frac{2\pi\hbar}{\frac{d\phi}{dx}} \tag{11.11}$$

Now, let's calculate the derivative:

$$\frac{d\lambda}{dx} \approx -\frac{2\pi\hbar}{\left(\frac{d\phi}{dx}\right)^2}\frac{d^2\phi}{dx^2} \tag{11.12}$$

So our condition $|d\lambda/dx| \ll 1$ is

$$\left| \frac{2\pi\hbar}{\left(\frac{d\phi}{dx}\right)^2} \frac{d^2\phi}{dx^2} \right| \ll 1 \quad \Longleftrightarrow \quad \left| \frac{1}{\hbar} \frac{d^2\phi}{dx^2} \right| \ll \frac{1}{2\pi} \frac{1}{\hbar^2} \left(\frac{d\phi}{dx} \right)^2 \qquad (11.13)$$

This tells us that the second term in the differential equation is small compared to the first. Let's define α to be the ratio of the two:

$$\alpha \equiv \frac{\frac{1}{\hbar} \frac{d^2\phi}{dx^2}}{\frac{1}{2\pi} \frac{1}{\hbar^2} \left(\frac{d\phi}{dx} \right)^2} \qquad (11.14)$$

Our approximation condition is therefore $|\alpha| \ll 1$.

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

$$\left(\frac{1}{\hbar} \frac{d\phi}{dx} \right)^2 - \frac{p^2}{\hbar^2} = 0 \quad \Longleftrightarrow \quad \frac{d\phi}{dx} = \pm p(x) \tag{11.15}$$

Integrating, we obtain

$$\phi(x) = \int^x dx' \, p(x') \tag{11.16}$$

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.

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

$$\phi(x) = \phi_0(x) + \alpha\,\phi_1(x) \tag{11.17}$$

and, similarly,

$$\frac{d}{dx}\,\phi(x) = \frac{d}{dx}\,\phi_0(x) + \alpha\,\frac{d}{dx}\,\phi_1(x) \tag{11.18}$$

$$\frac{d^2}{dx^2}\,\phi(x) = \frac{d^2}{dx^2}\,\phi_0(x) + \alpha\,\frac{d^2}{dx^2}\,\phi_1(x) \tag{11.19}$$

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:

$$\alpha \equiv \frac{\frac{1}{\hbar} \frac{d^2\phi}{dx^2}}{\frac{1}{2\pi} \frac{1}{\hbar^2} \left(\frac{d\phi}{dx}\right)^2} = \frac{\frac{1}{\hbar} \left(\frac{d^2\phi_0}{dx^2} + \alpha \frac{d^2\phi_1}{dx^2}\right)}{\frac{1}{2\pi} \frac{1}{\hbar^2} \left(\frac{d\phi_0}{dx} + \alpha \frac{d\phi_1}{dx}\right)^2} \approx \frac{\frac{1}{\hbar} \frac{d^2\phi_0}{dx^2}}{\frac{1}{2\pi} \frac{1}{\hbar^2} \left(\frac{d\phi_0}{dx}\right)^2} \tag{11.20}$$

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:

$$\left(\frac{\frac{d\phi_0}{dx} + \alpha \frac{d\phi_1}{dx}}{\hbar}\right)^2 - \frac{i}{\hbar} \frac{\frac{d^2\phi_0}{dx^2} + \alpha \frac{d^2\phi_1}{dx^2}}{\hbar} - \frac{p^2(x)}{\hbar^2} = 0 \tag{11.21}$$

$$\left(\frac{1}{\hbar} \frac{d\phi_0}{dx}\right)^2 + \frac{2\alpha}{\hbar^2} \frac{d\phi_0}{dx} \frac{d\phi_1}{dx} - \frac{i}{\hbar} \frac{d^2\phi_0}{dx^2} - \frac{p^2(x)}{\hbar^2} = 0 \tag{11.22}$$

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.

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, the coefficient of each power of α must vanish. We thus obtain two equations:

$$\left(\frac{1}{\hbar}\frac{d\phi_0}{dx}\right)^2 - \frac{p^2}{\hbar^2} = 0 \qquad \frac{2\alpha}{\hbar^2}\frac{d\phi_0}{dx}\frac{d\phi_1}{dx} = \frac{i}{\hbar}\frac{d^2\phi_0}{dx^2} \qquad (11.23)$$

The first equation is the same one we found when we made the approximation $\alpha = 0$; it yields

$$\frac{d\phi_0}{dx} = \pm p(x) \quad \Longleftrightarrow \quad \phi_0 = \pm \int^x dx' p(x') \qquad (11.24)$$

We see how our previous solution was indeed too gross an approximation: it consisted of setting $\alpha = 0$, neglecting any variation in the potential.

The second equation is

$$\frac{2\,\alpha}{\hbar^2}\,\frac{d\phi_0}{dx}\,\frac{d\phi_1}{dx} = \frac{i}{\hbar}\,\frac{d^2\phi_0}{dx^2} \tag{11.25}$$

$$-\frac{2\,i\,\alpha}{\hbar}\,\frac{d\phi_1}{dx} = \frac{\frac{d^2\phi_0}{dx^2}}{\frac{d\phi_0}{dx}} \tag{11.26}$$

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):

$$-\frac{2\,i\,\alpha}{\hbar}\,\frac{d\phi_1}{dx} = \frac{\frac{dp}{dx}}{p} \tag{11.27}$$

$$-\frac{2\,i\,\alpha}{\hbar}\phi_1 + c = \ln p \tag{11.28}$$

$$\frac{i}{\hbar}\,\alpha\,\phi_1 = -\frac{1}{2}\,\ln\frac{d\phi_0}{dx} + C = \ln\frac{1}{\sqrt{p}} + C \tag{11.29}$$

So, we have, to first order in α

$$\phi(x) \approx \phi_0(x) + \alpha \, \phi_1(x) = \pm \int^x dx' p(x') + \frac{\hbar}{i} \ln \frac{1}{\sqrt{p(x)}} + C \qquad (11.30)$$

$$\psi(x) = A \, e^{\frac{i}{\hbar} \phi(x)} = \frac{A}{\sqrt{p(x)}} \, \exp\left(\pm \frac{i}{\hbar} \int^x dx' p(x') \right) \qquad (11.31)$$

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.

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

$$\lambda(x) \approx \frac{2\pi\hbar}{\frac{d\phi}{dx}} \tag{11.32}$$

$$\frac{d\lambda}{dx} \approx \frac{\frac{1}{\hbar}\frac{d^2\phi}{dx^2}}{\frac{1}{2\pi\hbar}\left(\frac{d\phi}{dx}\right)^2} = \alpha \tag{11.33}$$

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.

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

$$V(x) = E - \frac{p^2(x)}{2\,m} = E - \frac{1}{2\,m}\left(\frac{d\phi_0}{dx}\right)^2 \qquad \frac{dV}{dx} = -\frac{1}{m}\frac{d\phi_0}{dx}\frac{d^2\phi_0}{dx^2} \qquad (11.34)$$

$$\alpha = \frac{\frac{1}{\hbar}\frac{d^2\phi_0}{dx^2}}{\frac{1}{2\pi}\frac{1}{\hbar^2}\left(\frac{d\phi_0}{dx}\right)^2} = -\frac{\frac{m}{\hbar}\frac{dV}{dx}}{\frac{1}{2\pi}\frac{1}{\hbar^2}\left(\frac{d\phi_0}{dx}\right)^3} = -\frac{\frac{dV}{dx}}{\frac{p^3(x)}{2\pi\hbar m}} = -\frac{\frac{dV}{dx}\frac{2\pi\hbar}{p}}{2\,T} \qquad (11.35)$$

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.

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, $T \to 0$, but the numerator is nonzero (in order for there to be a turning point!), so the approximation simply fails. This is equivalent to $p(x) \to 0$, 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:

$$J = -\frac{i}{2}\frac{\hbar}{m}\left(\psi^*(x)\frac{\partial}{\partial x}\psi(x) - \psi(x)\frac{\partial}{\partial x}\psi^*(x)\right) \tag{11.36}$$

$$= -\frac{i}{2}\frac{\hbar}{m}\left(\pm 2\frac{i}{\hbar}p(x)\right)\psi^*(x)\,\psi(x) \tag{11.37}$$

$$= P(x)\,v(x) \tag{11.38}$$

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.

Relation of the WKB Approximation to Classical Mechanics

The phase factor in the WKB approximation is

$$\phi_r(x) = \int^x dx' \, p(x') \tag{11.39}$$

(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

$$\phi_r(x_{cl}(t)) = \int^t dt \, \frac{dx_{cl}}{dt} \, p(x_{cl}(t)) \tag{11.40}$$

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

$$H = \frac{dx}{dt} \, p - L \tag{11.41}$$

We may then rewrite

$$\phi_r(x_{cl}(t)) = \int^t dt\, L_{cl} + \int^t dt\, H_{cl} \tag{11.42}$$

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

$$\phi_r(x_{cl}(t)) = \int^t dt\, L_{cl} + E\, t \tag{11.43}$$

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.

Thus, the wavefunction is

$$\psi(x) = \frac{A}{\sqrt{p(x)}} \, e^{\frac{i}{\hbar} \phi_r(x)} = \frac{A}{\sqrt{p(x_{cl}(t))}} \, \exp\left(\frac{i}{\hbar} \left[S_{cl}(x) + E \, t \right] \right) \tag{11.44}$$

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.

Lecture 29:
Semiclassical (WKB) Approximation

Revision Date: 2008/01/20

Derivation for Bound States

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

$$\psi(x) = \frac{A}{\sqrt{\kappa(x)}}\, \exp\left(\pm\frac{1}{\hbar}\int^x dx'\,\kappa(x')\right) \tag{11.45}$$

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.

Derivation for Bound States (cont.)

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.

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:

$$\psi_I(x) = \frac{A}{\sqrt{\kappa(x)}} \exp\left(\frac{1}{\hbar} \int_{x_1}^{x} dx' \, \kappa(x')\right) \tag{11.46}$$

$$\psi_{III}(x) = \frac{D}{\sqrt{\kappa(x)}} \exp\left(-\frac{1}{\hbar} \int_{x_2}^{x} dx' \, \kappa(x')\right) \tag{11.47}$$

$$\psi_{II,L}(x) = \frac{B}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar} \int_{x_1}^{x} dx' \, p(x') + \delta_L\right) \tag{11.48}$$

$$\psi_{II,R}(x) = \frac{C}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar} \int_{x}^{x_2} dx' \, p(x') + \delta_R\right) \tag{11.49}$$

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.

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

$$V(x) = \left\{ \begin{array}{ll} E - (x - x_1) F_1 & x \approx x_1 \\ E + (x - x_2) F_2 & x \approx x_2 \end{array} \right. \tag{11.50}$$

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

$$\frac{d^2\psi}{dx^2} + \frac{2\,m\,F_1}{\hbar^2}\,(x - x_1)\,\psi = 0 \qquad \frac{d^2\psi}{dx^2} - \frac{2\,m\,F_2}{\hbar^2}\,(x - x_1)\,\psi = 0 \tag{11.51}$$

These equations are of the same form,

$$\frac{d^2\psi}{dy^2} - y\,\psi = 0 \qquad \text{with} \quad y = \left\{ \begin{array}{ll} -\left(\frac{2\,m\,F_1}{\hbar^2}\right)^{1/3}(x - x_1) & x \approx x_1 \\ \left(\frac{2\,m\,F_2}{\hbar^2}\right)^{1/3}(x - x_2) & x \approx x_2 \end{array} \right. \tag{11.52}$$

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.

The solutions are termed Airy functions. They are given by

$$Ai(y) = a f(y) - b g(y) \qquad Bi(y) = \sqrt{3}\,[a f(y) - b g(y)] \tag{11.53}$$

$$a = \frac{1}{3^{2/3}\,\Gamma\left(\frac{2}{3}\right)} \approx 0.3550 \qquad b = \frac{1}{3^{1/3}\,\Gamma\left(\frac{1}{3}\right)} \approx 0.2588 \tag{11.54}$$

$$f(y) = 1 + \frac{1}{3!}\,y^3 + \frac{1 \cdot 4}{6!}\,y^6 + \frac{1 \cdot 4 \cdot 7}{9!}\,y^9 + \cdots \tag{11.55}$$

$$g(y) = y + \frac{2}{4!}\,y^4 + \frac{2 \cdot 5}{7!}\,y^7 + \frac{2 \cdot 5 \cdot 8}{10!}\,y^{10} + \cdots \tag{11.56}$$

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

$$Ai(y) \stackrel{y \to +\infty}{\longrightarrow} \frac{1}{2\sqrt{\pi}\, y^{1/4}} \exp\left(-\frac{2}{3}\, y^{3/2}\right) \tag{11.57}$$

$$Bi(y) \stackrel{y \to +\infty}{\longrightarrow} \frac{1}{\sqrt{\pi}\, y^{1/4}} \exp\left(\frac{2}{3}\, y^{3/2}\right) \tag{11.58}$$

$$Ai(y) \stackrel{y \to -\infty}{\longrightarrow} \frac{1}{\sqrt{\pi}\, (-y)^{1/4}} \sin\left(\frac{2}{3}\, (-y)^{3/2} + \frac{\pi}{4}\right) \tag{11.59}$$

$$Bi(y) \stackrel{y \to -\infty}{\longrightarrow} \frac{1}{\sqrt{\pi}\, (-y)^{1/4}} \cos\left(\frac{2}{3}\, (-y)^{3/2} + \frac{\pi}{4}\right) \tag{11.60}$$

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.

Let us do the matching near x_1. We need to recognize the following

$$p^2(x) \overset{x \to x_1^+}{=} 2\,m\,(E - V(x \to x_1^+)) = 2\,m\,F_1\,(x - x_1) = -(2\,m\,F_1\,\hbar)^{2/3}\,y \tag{11.61}$$

$$\kappa^2(x) \overset{x \to x_1^-}{=} 2\,m\,\left(V(x \to x_1^-) - E\right) = (2\,m\,F_1\,\hbar)^{2/3}\,y \tag{11.62}$$

$$2\,m\,F_1\,dx = -(2\,m\,F_1\,\hbar)^{2/3}\,dy \tag{11.63}$$

(Recall that y carries the opposite sign of $x - x_1$.) These forms let us explicitly do the WKB wavefunction argument integrals:

$$\frac{1}{\hbar} \int_{x_1}^{x} dx'\,\kappa(x') = -\int_0^y dy'\,\sqrt{y'} = -\frac{2}{3}\,y^{3/2} \qquad x < x_1, y > 0 \tag{11.64}$$

$$\frac{1}{\hbar} \int_{x_1}^{x} dx'\,p(x') = -\int_0^y dy'\,\sqrt{-y'} = \frac{2}{3}\,(-y)^{3/2} \qquad x > x_1, y < 0 \tag{11.65}$$

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.

So the WKB wavefunctions near x_1 become, when written in terms of y,

$$\psi_I(y) = \frac{1}{y^{1/4}} \, \exp\left(-\frac{2}{3}\, y^{3/2}\right) \qquad\qquad y > 0 \; (x < x_1) \qquad (11.66)$$

$$\psi_{II,L}(y) = \frac{B}{(-y)^{1/4}} \, \sin\left(\frac{2}{3}(-y)^{3/2} + \delta_L\right) \qquad y < 0 \; (x > x_1) \qquad (11.67)$$

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 $\delta_L = \pi/4$. (Note that the $\frac{1}{\sqrt{\pi}}$ factor in the Airy function asymptotic forms cancels away in connecting $\psi_I(y)$ to $\psi_{II,L}(y)$.)

Repeating the above procedure at $x \approx x_2$ will yield $C = 2D$ and $\delta_R = \pi/4$.

The remaining constant D is determined by matching the $\psi_{II,L}$ and $\psi_{II,R}$ solutions; explicitly:

$$\psi_{II,L}(x) = \psi_{II,R}(x) \qquad (11.68)$$

$$\frac{2}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_{x_1}^{x} dx'\, p(x') + \frac{\pi}{4}\right) = \frac{2D}{\sqrt{p(x)}} \sin\left(\frac{1}{\hbar}\int_{x}^{x_2} dx'\, p(x') + \frac{\pi}{4}\right) \quad (11.69)$$

Since D must be a constant, the equality can only hold if

$$\frac{1}{\hbar}\int_{x_1}^{x} dx'\, p(x') + \frac{\pi}{4} = -\left(\frac{1}{\hbar}\int_{x}^{x_2} dx'\, p(x') + \frac{\pi}{4}\right) + n\pi \qquad (11.70)$$

where n is to be determined. 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. Combining the two sides, we obtain

$$\int_{x_1}^{x_2} dx\, p(x) = \left(n\pi - \frac{\pi}{2}\right)\hbar = \left(n - \frac{1}{2}\right)\frac{h}{2} \qquad (11.71)$$

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.

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

$$\oint dx \, p(x) = \left(n + \frac{1}{2} \right) h \tag{11.72}$$

(The sign change on the $1/2$ term is absorbed by n).

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

$$\oint \frac{dx}{\lambda(x)} = n + \frac{1}{2} \tag{11.73}$$

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 12
Variational Method

Lecture 30:
Variational Method

Revision Date: 2008/01/16

Derivation

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.

Variational Method: Derivation

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,

$$E[\psi] = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\int dx \, \psi^*(x) H \psi(x)}{\int dx \, \psi^*(x)\psi(x)} \tag{12.1}$$

(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:

$$E + \delta E = \frac{\int dx \left[\psi(x) + \delta\psi(x)\right]^* H \left[\psi(x) + \delta\psi(x)\right]}{\int dx \left[\psi(x) + \delta\psi(x)\right]^* \left[\psi(x) + \delta\psi(x)\right]} \tag{12.2}$$

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

$$[E + \delta E] \left[\int dx \left[\psi(x) + \delta\psi(x)\right]^* \left[\psi(x) + \delta\psi(x)\right] \right] \tag{12.3}$$

$$= \int \left[\psi(x) + \delta\psi(x)\right]^* H \left[\psi(x) + \delta\psi(x)\right]$$

$$[E + \delta E] \left[\int dx \, \psi^*(x)\psi(x) + \int dx \, \delta\psi^*(x)\,\psi(x) + \int dx \, \psi^*(x)\,\delta\psi(x) \right] \tag{12.4}$$

$$= \int dx \, \psi^*(x) H \psi(x) + \int dx \, \delta\psi^*(x) \, H \psi(x) + \int dx \, \psi^*(x) H \, \delta\psi(x)$$

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.

$$E\left[\int dx\,\delta\psi^*(x)\,\psi(x) + \int dx\,\psi^*(x)\,\delta\psi(x)\right] = \int \delta\psi^*(x) H\,\psi(x) + \int \psi^*(x) H\,\delta\psi(x) \tag{12.5}$$

$$\int dx\,\delta\psi^*(x)\,(H-E)\,\psi(x) + \int dx\,\psi^*(x)\,(H-E)\,\delta\psi(x) = 0 \tag{12.6}$$

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

$$\int dx\,\chi(x)\,(H-E)\,\psi(x) = -\int dx\,\psi^*(x)\,(H-E)\,\chi(x) \tag{12.7}$$

Next, suppose $\delta\psi$ is completely imaginary, so $\delta\psi(x) = i\,\chi(x)$ where $\chi(x)$ is again real. This gives

$$\int dx\,\chi(x)\,(H-E)\,\psi(x) = \int dx\,\psi^*(x)\,(H-E)\,\chi(x) \tag{12.8}$$

where we have divided out by i.

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

$$\int dx\, \delta\psi^*(x)\, (H - E)\, \psi(x) = 0 \tag{12.9}$$

$$\int dx\, \psi^*(x)\, (H - E)\, \delta\psi(x) = 0 \tag{12.10}$$

Since $\delta\psi$ is arbitrary in each equation, the integrands must vanish. The first equation yields

$$(H - E)\, \psi(x) = 0 \tag{12.11}$$

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.

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

$$E'[\psi] = \int dx\, \psi^*(x)\, H\, \psi(x) - \lambda \left[\int dx\, \psi^*(x)\, \psi(x) - 1 \right] \tag{12.12}$$

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.

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:

$$|\delta\psi\,\rangle = \sum_n c_n\,|\psi_n\,\rangle \tag{12.13}$$

Now, let's calculate the energy of the wavefunction with the variation:

$$E[\psi_m + \delta\psi] = \frac{(\langle\psi_m\,| + \langle\delta\psi\,|)\,H\,(|\psi_m\,\rangle + |\delta\psi\,\rangle)}{(\langle\psi_m\,| + \langle\delta\psi\,|)\,(|\psi_m\,\rangle + |\delta\psi\,\rangle)} \tag{12.14}$$

$$= \frac{E_m + E_m\,(\langle\psi_m\,|\delta\psi\,\rangle + \langle\delta\psi\,|\psi_m\,\rangle)}{1 + (\langle\psi_m\,|\delta\psi\,\rangle + \langle\delta\psi\,|\psi_m\,\rangle)} \tag{12.15}$$

$$= E_m\left(\frac{1 + c_m + c_m^*}{1 + c_m + c_m^*}\right) = E_m \tag{12.16}$$

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 $\delta\psi$.

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:

$$E[\psi] = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2} \tag{12.17}$$

Subtract off the ground state energy E_0

$$E[\psi] - E_0 = \frac{\sum_n |c_n|^2 \, (E_n - E_0)}{\sum_n |c_n|^2} \tag{12.18}$$

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

$$E[\psi] \geq E_0 \tag{12.19}$$

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

$$\langle \phi_n | \psi \rangle = 0 \quad \text{for} \quad n < m \tag{12.20}$$

This condition can be met for any trial wavefunction $\psi(x)$ via Gram-Schmidt orthogonalization (Section 3.3). We then calculate the energy functional

$$E[\psi] = \frac{\sum_{n=m}^{\infty} |c_n|^2 E_n}{\sum_{n=0}^{\infty} |c_n|^2} \tag{12.21}$$

because the terms with $n < m$ vanish due to $c_n = \langle \phi_n | \psi \rangle = 0$. We then subtract E_m:

$$E[\psi] - E_m = \frac{\sum_{n=m}^{\infty} |c_n|^2 (E_n - E_m)}{\sum_{n=0}^{\infty} |c_n|^2} \tag{12.22}$$

$E_n \geq E_m$ for $n > m$, so we are assured the right side is nonnegative, yielding

$$E[\psi] \geq E_m \quad \text{when} \quad \langle \phi_n | \psi \rangle = 0 \quad \text{for} \quad n < m \tag{12.23}$$

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

$$\delta_0 = 1 - |\langle \psi_0 | \phi_0 \rangle|^2 \tag{12.24}$$

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

$$E[\psi_1] \geq E_1 - \delta_0 \left(E_1 - E_0 \right) \tag{12.25}$$

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.

Applications

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 12.1: 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

$$\psi(x; c) = \begin{cases} \left(\left(\frac{L}{2} \right)^2 - x^2 \right) (1 + c\,x^2) & |x| \leq \frac{L}{2} \\ 0 & |x| > \frac{L}{2} \end{cases} \tag{12.26}$$

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

$$E[\psi] \to E[c] = \frac{3\,\hbar^2}{m\,L^2} \, \frac{11\left(\frac{L}{2}\right)^4 c^2 + 14\left(\frac{L}{2}\right)^2 c + 35}{\left(\frac{L}{2}\right)^4 c^2 + 6\left(\frac{L}{2}\right)^2 c + 21} \tag{12.27}$$

We then find the extrema of $E(c)$ with respect to c, $dE/dc = 0$, yielding

$$26\left(\frac{L}{2}\right)^4 c^2 + 196\left(\frac{L}{2}\right)^2 c + 42 = 0 \tag{12.28}$$

with solutions

$$c_1 = -\frac{0.221075}{\left(\frac{L}{2}\right)^2} \qquad c_2 = -\frac{7.31771}{\left(\frac{L}{2}\right)^2} \tag{12.29}$$

Applications (cont.)

We may then calculate E for these values, giving

$$E(c_1) = 4.93488\,\frac{\hbar^2}{m\,L^2} \qquad E(c_2) = 51.0652\,\frac{\hbar^2}{m\,L^2} \tag{12.30}$$

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

$$E_{n=1} = \frac{\hbar^2 \pi^2}{2\,m\,L^2} = 4.93480\,\frac{\hbar^2}{m\,L^2} \tag{12.31}$$

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.

Another interesting fact is that the second solution, using c_2, is a decent approximation to the second excited state, $n = 3$:

$$E_{n=3} = 9\, E_{n=1} = 44.4132 \, \frac{\hbar^2}{m\, L^2} \tag{12.32}$$

There is a general theorem on this point, the Hylleraas-Undheim Theorem, which essentially states that, if the trial wavefunction is 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.

Applications (cont.)

Example 12.2: 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 13
Rotations and Orbital Angular Momentum

Lecture 31:
Rotations and Orbital Angular Momentum in Two Dimensions

Revision Date: 2008/01/16

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 9, especially Section 9.3, so please review that material.

Rotations Transformations in Two Dimensions

Passive Classical Rotation Transformations in Two Dimensions

Recall that 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'):

$$a_x' = a_x \cos\theta + a_y \sin\theta \qquad a_y' = -a_x \sin\theta + a_y \cos\theta \qquad a_z' = z \quad (13.1)$$

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:

$$\begin{bmatrix} a_x' \\ a_y' \\ a_z' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & & 1 \end{bmatrix} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \equiv \mathbf{R}_{P,\theta\widehat{z}} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \quad (13.2)$$

where we use the $_P$ subscript to indicate a passive transformation (as we did in the QM case) and the $_{\theta\widehat{z}}$ subscript to indicate the rotation angle from the untransformed to the transformed system.

Rotations Transformations in Two Dimensions (cont.)

Active Classical Rotation Transformations in Two Dimensions

The classical analogue of an *active* coordinate transformation is to change the vector; that is, to change the representation of the vector in a given coordinate system. If we denote the new vector by $\vec{b}$, then the components of $\vec{b}$ are related to those of $\vec{a}$ by

$$
\left[
\begin{array}{c}
b_x \\
b_y \\
b_z
\end{array}
\right]
=
\left[
\begin{array}{ccc}
\cos\theta & -\sin\theta & 0 \\
\sin\theta & \cos\theta & 0 \\
0 & & 1
\end{array}
\right]
\left[
\begin{array}{c}
a_x \\
a_y \\
a_z
\end{array}
\right]
\equiv \mathbf{R}_{A,\theta\widehat{z}}
\left[
\begin{array}{c}
a_x \\
a_y \\
a_z
\end{array}
\right]
\tag{13.3}
$$

where both are being represented in the untransformed coordinate system. This transformation corresponds to physically rotating $\vec{a}$ by θ CCW about $\widehat{z}$.

Rotations Transformations in Two Dimensions (cont.)

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 $\vec{b}$, while the passive transformation rotates the coordinate system so that the representation of the vector $\vec{a}$ changes to $\vec{a}\,'$; but $\vec{a}$ and $\vec{a}\,'$ are physically the same vector.

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 $\widehat{z}$.

The mathematical difference between the passive and active transformations is just the change of sign of the $\sin\theta$ terms; that is $R_{P,\theta\widehat{z}} = R_{A,-\theta\widehat{z}}$. This sign flip tells us that, $\vec{a}\,'$, the representation of $\vec{a}$ in a coordinate system whose unit vectors are obtained by active rotation of the untransformed system's unit vectors by $\theta\widehat{z}$, is mathematically equivalent to the representation in the untransformed coordinate system of a vector $\vec{b}$ that has been obtained from $\vec{a}$ by active rotation by $-\theta\widehat{z}$. $\vec{a}\,'$ and $\vec{b}$ are different vectors because, even through the representations are mathematically equivalent, they are referenced to different coordinate systems (the transformed and untransformed systems).

It is convention to use $\mathbf{R}_{\theta\widehat{z}}$ for $\mathbf{R}_{A,\theta\widehat{z}}$ and to never use $\mathbf{R}_{P,\theta\widehat{z}}$. 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:

$$\mathbf{R}_{\delta\theta\widehat{z}} = \left[\begin{array}{ccc} \cos\delta\theta & -\sin\delta\theta & 0 \\ \sin\delta\theta & \cos\delta\theta & 0 \\ 0 & & 1 \end{array} \right] \approx \left[\begin{array}{ccc} 1 & -\delta\theta & 0 \\ \delta\theta & 1 & 0 \\ 0 & & 1 \end{array} \right] \qquad (13.4)$$

Rotations Transformations in Two Dimensions (cont.)

The generic relationship between a classical coordinate transformation and its generator is

$$\mathbf{T}_\epsilon = \mathbf{I} + \epsilon\,\mathbf{G} \tag{13.5}$$

Instead of relating Hermitian generators to unitary coordinate transformation operators, we must relate symmetric generators to orthogonal coordinate transformation operators. Thus, it makes sense to rewrite our infinitesimal rotation operators as

$$\mathbf{R}_{\delta\theta\widehat{z}} = \mathbf{I} + \delta\theta\,\mathbf{M}_z \qquad \mathbf{M}_z \equiv \left[\begin{array}{ccc} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right] \tag{13.6}$$

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 $\widehat{y}$.

We of course recover the finite classical rotation transformation by the appropriate infinite product, yielding an exponential:

$$\mathbf{R}_{\theta\widehat{z}} = \lim_{N\to\infty} \left(\mathbf{I} + \frac{\theta}{N}\,\mathbf{M}_z \right)^N = \exp\left(\theta\,\mathbf{M}_z\right) \tag{13.7}$$

We may evaluate the above using the fact

$$\mathbf{M}_z^2 = - \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right] \tag{13.8}$$

This yields

$$\mathbf{R}_{\theta\hat{z}} = \sum_{n=0}^{\infty} \frac{\theta^n}{n!} \mathbf{M}_z^n = \mathbf{I} + \theta\,\mathbf{M}_z + \sum_{n=1}^{\infty} \left(\frac{\theta^{2n}}{(2n)!}\mathbf{M}_z^{2n} + \frac{\theta^{2n+1}}{(2n+1)!}\mathbf{M}_z^{2n}\mathbf{M}_z \right) \quad (13.9)$$

$$= \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{array} \right] + \sum_{n=0}^{\infty} \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right] \left(\frac{\theta^{2n}(-1)^n}{(2n)!} + \frac{\theta^{2n+1}(-1)^n}{(2n+1)!}\mathbf{M}_z \right) \quad (13.10)$$

$$= \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{array} \right] + \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right] (\cos\theta + \sin\theta\,\mathbf{M}_z) \quad (13.11)$$

$$= \left[\begin{array}{ccc} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{array} \right] \quad (13.12)$$

as expected.

What is the significance of the $\mathbf{M}_z$ matrix? See:

$$-\vec{r}^T \mathbf{M}_z \vec{p} = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} \tag{13.13}$$

$$= x\, p_y - y\, p_x = l_z \tag{13.14}$$

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.

Quantum Mechanical Active Rotation Transformation in Two Dimensions

Now, let's consider the quantum mechanical version of the above. We want to obtain the quantum mechanical rotation transformation operator and its generator for a transformation by $\theta \widehat{z}$.

Let's recall in detail what we mean by the above. We consider the classical coordinate transformation described earlier in this section, where the $\widehat{x}\,'$ and $\widehat{y}\,'$ unit vectors are obtained by rotating the $\widehat{x}$ and $\widehat{y}$ unit vectors by θ CCW about $\widehat{z}$. There are position-basis kets in the untransformed system $\{|x, y\,\rangle\}$ and in the transformed system $\{|x', y'\,\rangle\}$. For a generic transformation, they are related by

$$|x' = u, y' = v\,\rangle = T(\theta \widehat{z})|x = u, y = v\,\rangle \qquad (13.15)$$

The first question to ask is: how are the transformed system position-basis kets decomposed in terms of the untransformed system position-basis kets? That is, we have a position-basis ket $|x' = u, y' = v\,\rangle$ that corresponds to a particle isolated at the position $x' = u$, $y' = v$. What position-basis ket does this correspond to in the untransformed basis?

Rotations Transformations in Two Dimensions (cont.)

We need our classical rotation transformation to determine this. The passive rotation transformation tells us that

$$\left[\begin{array}{c} b_x \\ b_y \end{array}\right] = \left[\begin{array}{cc} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{array}\right] \left[\begin{array}{c} a_x \\ a_y \end{array}\right] \tag{13.16}$$

(We drop the z dimension for brevity.) We take $b_x = u$, $b_y = v$, and invert the above:

$$\left[\begin{array}{c} a_x \\ a_y \end{array}\right] = \left[\begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array}\right] \left[\begin{array}{c} u \\ v \end{array}\right] = \left[\begin{array}{c} u\cos\theta - v\sin\theta \\ u\sin\theta + v\cos\theta \end{array}\right] \tag{13.17}$$

Therefore,

$$T(\theta\hat{z})|x = u, y = v\,\rangle = |x' = u, y' = v\,\rangle = |x = u\cos\theta - v\sin\theta, y = u\sin\theta + v\cos\theta\,\rangle \tag{13.18}$$

The next step is to use unitarity and the above to obtain the explicit form of the generator of $T(\theta \widehat{z})$ and from there to obtain the explicit form for $T(\theta \widehat{z})$. That is, we write an infinitesimal version of the rotation transformation operator in terms of a generator to be determined:

$$T(\delta\theta\widehat{z}) = I - \frac{i}{\hbar}\,\delta\theta\,G \tag{13.19}$$

Next, we calculate the position-basis representation of a transformed state in the untransformed position basis using the definition of the infinitesimal rotation operator in terms of the generator:

$$\langle x = u, y = v \,|\psi'\,\rangle = \langle x = u, y = v \,|T(\delta\theta\widehat{z})|\psi\rangle \tag{13.20}$$

$$= \langle x = u, y = v \,|\psi\rangle - \delta\theta\,\frac{i}{\hbar}\langle x = u, y = v \,|G|\psi\rangle \tag{13.21}$$

$$= \psi_q(x = u, y = v) - \delta\theta\,\frac{i}{\hbar}\langle x = u, y = v \,|G|\psi\rangle \tag{13.22}$$

where, as in Section 9, $\psi_q(x = u, y = v) \equiv \langle x = u, y = v \,|\psi\rangle$ is the representation of the untransformed state $|\psi\rangle$ in terms of the untransformed position basis $\{|x, y\rangle\}$.

Rotations Transformations in Two Dimensions (cont.)

Next, we evaluate the same expression but allowing T to act to the left:

$$\langle x = u, y = v \,|\, T(\delta\theta\hat{z})|\psi\,\rangle \tag{13.23}$$

$$= \left(T^{\dagger}(\delta\theta\hat{z})|x = u, y = v\,\rangle \right)^{\dagger} |\psi\,\rangle \tag{13.24}$$

$$= (|x = u + v\,\delta\theta, y = -u\,\delta\theta + v\,\rangle)^{\dagger} |\psi\,\rangle \tag{13.25}$$

$$= \langle x = u + v\,\delta\theta, y = -u\,\delta\theta + v \,|\psi\,\rangle \tag{13.26}$$

$$= \psi_q(x = u + v\,\delta\theta, y = -u\,\delta\theta + v) \tag{13.27}$$

$$= \psi_q(x = u, y = v) + v\,\delta\theta\frac{\partial}{\partial x}\,\psi_q(x = u, y = v) - u\,\delta\theta\frac{\partial}{\partial y}\,\psi_q(x = u, y = v) \tag{13.28}$$

$$= \psi_q(x = u, y = v) - \delta\theta\,\frac{i}{\hbar}\left(\langle x = u, y = v \,|X\,P_y|\psi\,\rangle - \langle x = u, y = v \,|Y\,P_x|\psi\,\rangle \right) \tag{13.29}$$

$$= \psi_q(x = u, y = v) - \delta\theta\,\frac{i}{\hbar}\langle x = u, y = v \,|\,(X\,P_y - Y\,P_x)\,|\psi\,\rangle \tag{13.30}$$

where, in the second step, the sign of θ arises from the fact that we are acting with $T^{\dagger}$, not T.

Equating the two expressions and noting that $\langle x = u, y = v \,|$ and $|\,\psi\,\rangle$ are arbitrary lets us conclude

$$G = X\,P_y - Y\,P_x \equiv L_z \tag{13.31}$$

You will recognize this as the quantum analogue of the z-axis angular momentum, $l_z = x\,p_y - y\,p_x$, hence we denote it above by L_z. It is Hermitian because each operator that enters is Hermitian and both products commute. From the above, we may conclude that the generic explicit form for the quantum mechanical rotation operator is

$$T(\theta\widehat{z}) = \exp\left(-\frac{i}{\hbar}\,\theta\,L_z\right) \tag{13.32}$$

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.

Rotations 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:

$$\frac{i}{\hbar}\,\theta\,\langle x, y\,|L_z|\psi\rangle = \frac{i}{\hbar}\,\theta\,\langle x, y\,|\,(X\,P_y - Y\,P_x)\,|\psi\rangle = \theta\left(x\,\frac{\partial}{\partial y} - y\,\frac{\partial}{\partial x}\right)\psi_q(x, y)$$

(13.33)

Now, we need to change variables to polar coordinates. The functional relationship between polar and cartesian coordinates is

$$\rho = \sqrt{x^2 + y^2} \qquad \phi = \arctan\frac{y}{x}$$

(13.34)

Hence

$$\frac{\partial\rho}{\partial x} = \frac{x}{\rho} \qquad \frac{\partial\rho}{\partial y} = \frac{y}{\rho} \qquad \frac{\partial\phi}{\partial x} = \frac{-y/x^2}{1 + (y/x)^2} = -\frac{y}{\rho^2} \qquad \frac{\partial\phi}{\partial y} = \frac{1/x}{1 + (y/x)^2} = \frac{x}{\rho^2}$$

(13.35)

Rotations Transformations in Two Dimensions (cont.)

The chain rule thus tells us

$$\frac{\partial}{\partial x} = \frac{\partial \rho}{\partial x} \frac{\partial}{\partial \rho} + \frac{\partial \phi}{\partial x} \frac{\partial}{\partial \phi} = \frac{x}{\rho} \frac{\partial}{\partial \rho} - \frac{y}{\rho^2} \frac{\partial}{\partial \phi} \tag{13.36}$$

$$\frac{\partial}{\partial y} = \frac{\partial \rho}{\partial y} \frac{\partial}{\partial \rho} + \frac{\partial \phi}{\partial y} \frac{\partial}{\partial \phi} = \frac{y}{\rho} \frac{\partial}{\partial \rho} + \frac{x}{\rho^2} \frac{\partial}{\partial \phi} \tag{13.37}$$

So, then,

$$\theta \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \psi_q(x, y) = \theta \left(\frac{\partial}{\partial \phi} \right) \psi_q(\rho, \phi) \tag{13.38}$$

where we simply rewrite ψ_q in terms of ρ and ϕ using $x = \rho \cos \phi$ and $y = \rho \sin \phi$. So,

$$\langle x, y | \frac{i}{\hbar} \theta L_z | \psi \rangle = \langle x, y | \frac{i}{\hbar} \theta (X P_y - Y P_x) | \psi \rangle \tag{13.39}$$

$$= \langle x, y | \theta \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) | \psi \rangle = \langle x, y | \theta \frac{\partial}{\partial \phi} | \psi \rangle \tag{13.40}$$

which looks like the action of the generator of a translation in the polar angle ϕ by an angle θ, as we expect.

Rotations Transformations in Two Dimensions (cont.)

The last thing we need to consider is the action of the rotation transformation on the standard operators X, Y, P_x, and P_Y. For any operator Q, a coordinate transformation of any kind defines a new operator $Q' = T Q T^{\dagger}$. What is the effect of the coordinate transformation on these specific operators? We could use the methods used in Example 9.5 to calculate these 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. We will need the relation

$$e^{-A} B e^{A} = B + [B, A] + \frac{1}{2!} \, [[B, A], A] + + \frac{1}{3!} \, [[[B, A], A], A] \cdots \tag{13.41}$$

(which we do not prove here). This relation will make use of the following important commutators:

$$[X, L_z] = [X, X P_y] - [X, Y P_x] = 0 - Y [X, P_x] = -i \hbar Y \tag{13.42}$$

$$[Y, L_z] = [Y, X P_y] - [Y, Y P_x] = X [Y, P_y] - 0 = i \hbar X \tag{13.43}$$

$$[P_x, L_z] = [P_x, X P_y] - [P_x, Y P_x] = [P_x, X] P_y - 0 = -i \hbar P_y \tag{13.44}$$

$$[P_y, L_z] = [P_y, X P_y] - [P_y, Y P_x] = 0 - [P_y, Y] P_x = i \hbar P_x \tag{13.45}$$

The evident cyclicity of the above relations will be written succinctly when we consider rotations in three dimensions.

With the above, the transformed operators are easily evaluated:

$$X' = T(\theta\widehat{z}) \, X \, T(-\theta\widehat{z}) = e^{-\frac{i}{\hbar}\,\theta\,L_z} \, X \, e^{\frac{i}{\hbar}\,\theta\,L_z} \tag{13.46}$$

$$= X + \frac{i}{\hbar}\,\theta[X, L_z] + \frac{1}{2!}\left(\frac{i}{\hbar}\,\theta\right)^2 [[X, L_z], L_z] + \frac{1}{3!}\left(\frac{i}{\hbar}\,\theta\right)^3 [[[X, L_z], L_z], L_z] + \cdots \tag{13.47}$$

$$= X + \theta\,Y + \frac{\theta^2}{2!}\,(-1)\,X + \frac{\theta^3}{3!}\,(-1)\,Y + \cdots \tag{13.48}$$

$$= X\left(1 - \frac{\theta^2}{2!} + \cdots\right) + Y\left(\theta - \frac{\theta^3}{3!} + \cdots\right) \tag{13.49}$$

$$= X\cos\theta + Y\sin\theta \tag{13.50}$$

$$Y' = -X\sin\theta + Y\cos\theta \tag{13.51}$$

$$P_x' = P_x\cos\theta + P_y\sin\theta \tag{13.52}$$

$$P_y' = -P_x\sin\theta + P_y\cos\theta \tag{13.53}$$

where the last three are evaluated in the same way as X'.

Lecture 32:
The Eigenvalue Problem of L_z in Two Dimensions
Rotations and Angular Momentum in Three Dimensions

Revision Date: 2008/01/16

The Eigenvalue Problem of L_z in Two Dimensions

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:

$$L_z |l_z\rangle = l_z |l_z\rangle \tag{13.54}$$

where l_z is an eigenvalue and $|l_z\rangle$ 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:

$$-i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\psi_{l_z}(x,y) = l_z\,\psi_{l_z}(x,y) \tag{13.55}$$

We also make use of the change of variables to (ρ, ϕ) to obtain a simpler equation:

$$-i\hbar\frac{\partial}{\partial\phi}\psi_{l_z}(\rho,\phi) = l_z\,\psi_{l_z}(\rho,\phi) \tag{13.56}$$

The solution is obvious,

$$\psi_{l_z}(\rho, \phi) = R(\rho)\, e^{\frac{i}{\hbar}\, \phi\, l_z} \tag{13.57}$$

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

$$\langle \psi_1 \,|L_z|\psi_2 \,\rangle = \langle \psi_2 \,|L_z|\psi_1 \,\rangle^* \tag{13.58}$$

which, written out in terms of the position-basis wavefunction, is

$$\int_0^\infty d\rho\, \rho \int_0^{2\pi} d\phi\, \psi_1^*(\rho, \phi) \left(-i\,\hbar\, \frac{\partial}{\partial \phi}\right) \psi_2(\rho, \phi)$$
$$= \left[\int_0^\infty d\rho\, \rho \int_0^{2\pi} d\phi\, \psi_2^*(\rho, \phi) \left(-i\,\hbar\, \frac{\partial}{\partial \phi}\right) \psi_1(\rho, \phi)\right]^* \tag{13.59}$$

To obtain a condition on ψ_1 and ψ_2, we integrate the right side by parts, yielding

$$\text{RHS} = i\hbar \int_0^\infty d\rho\, \rho \left[\psi_2(\rho,\phi)\,\psi_1^*(\rho,\phi)|_0^{2\pi} - \int_0^{2\pi} d\phi\, \psi_1^*(\rho,\phi) \left(\frac{\partial}{\partial\phi} \right) \psi_2(\rho,\phi) \right] \tag{13.60}$$

$$= i\hbar \int_0^\infty d\rho\, \rho\, \psi_2(\rho,\phi)\,\psi_1^*(\rho,\phi)|_0^{2\pi} + \text{LHS} \tag{13.61}$$

We require RHS = LHS for any ψ_1, ψ_2 (not just eigenfunctions), including any possible radial dependence, so we must have

$$\psi(\rho, 0) = \psi(\rho, 2\pi) \tag{13.62}$$

at any ρ for any ψ. If we impose the constrain on the eigenfunctions, we have

$$1 = e^{2\pi \frac{i}{\hbar} l_z} \tag{13.63}$$

which implies

$$l_z = m\hbar \qquad m = 0, \pm 1, \pm 2, \ldots \tag{13.64}$$

One could obtain the same result by requiring that the eigenfunctions be single-valued,

$$\psi(\rho, \phi + 2\pi) = \psi(\rho, \phi) \tag{13.65}$$

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,

$$|\psi(\rho, \phi + 2\pi)|^2 = |\psi(\rho, \phi)|^2 \tag{13.66}$$

$$\left| R(\rho)\, e^{\frac{i}{\hbar} l_z \phi}\, e^{\frac{i}{\hbar} 2\pi l_z} \right|^2 = \left| R(\rho)\, e^{\frac{i}{\hbar} l_z \phi} \right|^2 \tag{13.67}$$

This results in the requirement $\left| e^{\frac{i}{\hbar} 2\pi l_z} \right| = 1$, 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.

It turns out that, if one considers superposition states of the form

$$\psi(\rho, \phi) = A(\rho)\, e^{\frac{i}{\hbar}\, l_z\, \phi} + B(\rho)\, e^{\frac{i}{\hbar}\, l_z'\, \phi} \tag{13.68}$$

and requires that their probability density always be single-valued, then one can obtain the condition $l_z - l_z' = m\,\hbar$, $m = 0, \pm 1, \pm 2, \ldots$. 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.

We shall take as our normalized azimuthal eigenfunctions

$$\Phi_m(\phi) = \frac{1}{\sqrt{2\pi}} \, e^{i m \phi} \qquad m = 0, \pm 1, \pm 2, \dots \qquad (13.69)$$

They obey the orthonormality condition

$$\int_0^{2\pi} d\phi \, \Phi_m^*(\phi) \, \Phi_{m'}(\phi) = \delta_{mm'}, \qquad (13.70)$$

The full eigenfunctions are of the form

$$\psi(\rho, \phi) = R(\rho) \, \Phi_m(\phi) \qquad (13.71)$$

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:

$$[P_x^2 + P_y^2, L_z] = P_x P_x L_z - L_z P_x P_x + P_y P_y L_z - L_z P_y P_y \tag{13.72}$$

$$= P_x [P_x, L_z] + [P_x, L_z] P_x + P_y [P_y, L_z] + [P_y, L_z] P_y \tag{13.73}$$

$$= -i\hbar P_x P_y - i\hbar P_y P_x + i\hbar P_y P_x + i\hbar P_x P_y = 0 \tag{13.74}$$

Second, it is obvious from the form of L_z when projected onto the position basis and
written in polar coordinates that $[L_z, V(\rho)] = 0$. Therefore, $[H, L_z] = 0$.

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

$$\left[-\frac{\hbar^2}{2\mu} \left(\frac{\partial^2}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial\phi^2} \right) + V(\rho) \right] \psi_E(\rho, \phi) = E\,\psi_E(\rho, \phi) \qquad (13.75)$$

(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 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.

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.

Inserting the form $\psi_E(\rho, \phi) = R_{E,m}(\rho)\, \Phi_m(\phi)$ into the Hamiltonian's eigenvector-eigenvalue equation yields

$$\left[-\frac{\hbar^2}{2\,\mu} \left(\frac{\partial^2}{\partial\rho^2} + \frac{1}{\rho}\,\frac{\partial}{\partial\rho} - \frac{m^2}{\rho^2} \right) + V(\rho) \right] R_{E,m}(\rho)\, \Phi_m(\phi) = E\, R_{E,m}(\rho)\, \Phi_m(\phi) \quad (13.76)$$

$\Phi_m(\phi)$ 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

$$\left[-\frac{\hbar^2}{2\,\mu} \left(\frac{d^2}{d\rho^2} + \frac{1}{\rho}\,\frac{d}{d\rho} - \frac{m^2}{\rho^2} \right) + V(\rho) \right] R_{E,m}(\rho) = E\, R_{E,m}(\rho) \quad (13.77)$$

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 $V(\rho)$.

Rotations and Angular Momentum in Three Dimensions

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:

$$\mathbf{R}_{\theta\widehat{x}} = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{array}\right] \qquad \mathbf{R}_{\theta\widehat{y}} = \left[\begin{array}{ccc} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{array}\right] \qquad (13.78)$$

$$\mathbf{R}_{\delta\theta\widehat{x}} = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & -\delta\theta \\ 0 & \delta\theta & 1 \end{array}\right] = \mathbf{I} + \delta\theta\,\mathbf{M}_x \qquad \mathbf{M}_x \equiv \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{array}\right] \qquad (13.79)$$

$$\mathbf{R}_{\delta\theta\widehat{y}} = \left[\begin{array}{ccc} 1 & 0 & \delta\theta \\ 0 & 1 & 0 \\ -\delta\theta & 0 & 1 \end{array}\right] = \mathbf{I} + \delta\theta\,\mathbf{M}_y \qquad \mathbf{M}_y \equiv \left[\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{array}\right] \qquad (13.80)$$

Infinitesimal rotations about different axes commute because they are infinitesimal. For example,

$$\mathbf{R}_{\delta\theta_x\hat{x}}\mathbf{R}_{\delta\theta_y\hat{y}} = (\mathbf{I} + \delta\theta_x\mathbf{M}_x)(\mathbf{I} + \delta\theta_y\mathbf{M}_y) = \mathbf{I} + \delta\theta_x\mathbf{M}_x + \delta\theta_y\mathbf{M}_y + O(\delta\theta)^2 \qquad (13.81)$$

$$\mathbf{R}_{\delta\theta_y\hat{y}}\mathbf{R}_{\delta\theta_x\hat{x}} = (\mathbf{I} + \delta\theta_y\mathbf{M}_y)(\mathbf{I} + \delta\theta_x\mathbf{M}_x) = \mathbf{I} + \delta\theta_x\mathbf{M}_x + \delta\theta_y\mathbf{M}_y + O(\delta\theta)^2 \qquad (13.82)$$

$$\approx \mathbf{R}_{\delta\theta_x\hat{x}}\mathbf{R}_{\delta\theta_y\hat{y}} \qquad (13.83)$$

The generic form for an infinitesimal rotation is therefore

$$\mathbf{R}_{\delta\vec{\theta}} = \mathbf{I} + \delta\theta_x\mathbf{M}_x + \delta\theta_y\mathbf{M}_y + \delta\theta_z\mathbf{M}_z \equiv \mathbf{I} + \delta\vec{\theta}\cdot\vec{\mathbf{M}} \qquad (13.84)$$

$$\text{with} \quad \vec{\mathbf{M}} = \hat{x}\,\mathbf{M}_x + \hat{y}\,\mathbf{M}_y + \hat{z}\,\mathbf{M}_z \qquad (13.85)$$

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.

It follows from the above that finite rotations may be written in the form

$$\mathbf{R}_{\vec{\theta}} = \exp(\vec{\theta} \cdot \vec{\mathbf{M}}) = \exp(\theta_x \mathbf{M}_x + \theta_y \mathbf{M}_y + \theta_z \mathbf{M}_z) \tag{13.86}$$

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.

We make a few more useful points about the $\vec{\mathbf{M}}$ matrices. First, an easy-to-remember form for them is

$$(\mathbf{M}_i)_{jk} = -\epsilon_{ijk} \tag{13.87}$$

where $i = 1, 2, 3$ corresponds to $i = x, y, z$ and where ϵ_{ijk} 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:

$$\mathbf{M}_x^2 = - \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right] \qquad \mathbf{M}_y^2 = - \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{array} \right] \qquad \mathbf{M}_z^2 = - \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right]$$

(13.88)

Therefore,

$$|\vec{\mathbf{M}}|^2 = \vec{\mathbf{M}} \cdot \vec{\mathbf{M}} = \mathbf{M}_x^2 + \mathbf{M}_y^2 + \mathbf{M}_z^2 = -2\,\mathbf{I}$$

(13.89)

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

$$[\mathbf{M}_i, \mathbf{M}_j] = \epsilon_{ijk}\mathbf{M}_k \qquad \Longleftrightarrow \qquad \vec{\mathbf{M}} \times \vec{\mathbf{M}} = \vec{\mathbf{M}} \tag{13.90}$$

which 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

$$[\mathbf{M}_i, \vec{\mathbf{M}} \cdot \vec{\mathbf{M}}] = 0 \tag{13.91}$$

Finally, just as $\vec{r}^T \mathbf{M}_z \vec{p} = -l_z$, we have in general

$$\vec{l} = -\vec{r}^T \vec{\mathbf{M}}\, \vec{p} \tag{13.92}$$

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

$$L_x = Y P_z - Z P_y \qquad L_y = Z P_x - X P_z \qquad L_z = X P_y - Y P_x \qquad (13.93)$$

$$\vec{L} = \widehat{x} L_x + \widehat{y} L_y + \widehat{z} L_z \qquad (13.94)$$

The components satisfy the commutation relation

$$[L_i, L_j] = \epsilon_{ijk}\, i\, \hbar\, L_k \qquad \Longleftrightarrow \qquad \vec{L} \times \vec{L} = i\, \hbar\, \vec{L} \qquad (13.95)$$

just as the $\mathbf{M}_i$ 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

$$T(\vec{\theta}) = \exp\left(-\frac{i}{\hbar}\, \vec{\theta} \cdot \vec{L}\right) \qquad (13.96)$$

Additionally, one can show

$$[R_i, L_j] = \epsilon_{ijk}\, i\, \hbar\, R_k \qquad [P_i, L_j] = \epsilon_{ijk}\, i\, \hbar\, P_k \tag{13.97}$$

where R_i are the position component operators X, Y, and Z.

Just as we calculated $\vec{\mathbf{M}} \cdot \vec{\mathbf{M}}$, we may also calculate L^2,

$$L^2 = L_x^2 + L_y^2 + L_z^2 \tag{13.98}$$

One may verify that

$$[L_i, L^2] = 0 \tag{13.99}$$

which recalls a similar property of $\vec{\mathbf{M}} \cdot \vec{\mathbf{M}}$. So, while no two of the $\{L_i\}$ are simultaneously diagonalizable, one may simultaneously diagonalize any one of the $\{L_i\}$ and the L^2 operator.

Lecture 33:
The Eigenvector-Eigenvalue Problem of L_z and L^2

Revision Date: 2008/01/20

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 straightforward 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 make obtaining the full position-basis representations more convenient.

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

$$r = \sqrt{x^2 + y^2 + z^2} \qquad \theta = \arccos \frac{z}{r} \qquad \phi = \arctan \frac{y}{x} \tag{13.100}$$

We can convert the cartesian representations of the L_i to spherical coordinates via chain rule as we did in two dimensions, though it is more tedious here. The result is

$$\langle r, \theta, \phi \, | L_x | \psi \, \rangle = i \, \hbar \left(\sin \phi \, \frac{\partial}{\partial \theta} + \cot \theta \, \cos \phi \, \frac{\partial}{\partial \phi} \right) \psi_q(r, \theta, \phi) \tag{13.101}$$

$$\langle r, \theta, \phi \, | L_y | \psi \, \rangle = i \, \hbar \left(-\cos \phi \, \frac{\partial}{\partial \theta} + \cot \theta \, \sin \phi \, \frac{\partial}{\partial \phi} \right) \psi_q(r, \theta, \phi) \tag{13.102}$$

$$\langle r, \theta, \phi \, | L_z | \psi \, \rangle = -i \, \hbar \, \frac{\partial}{\partial \phi} \, \psi_q(r, \theta, \phi) \tag{13.103}$$

We will also need the L^2 operator, which is straightforward to calculate from the above:

$$\langle r, \theta, \phi \,|L^2|\psi\rangle = -\hbar^2 \left(\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \sin\theta \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} \right) \psi_q(r, \theta, \phi) \qquad (13.104)$$

Note that the first $\frac{\partial}{\partial\theta}$ acts on everything to its right including the $\sin\theta$ factor.

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

$$\psi_{\alpha,m}(r, \theta, \phi) = R(r)\Theta_{\alpha,m}(\theta)\,\Phi_m(\phi) \qquad (13.105)$$

$$\Phi_m(\phi) = \frac{e^{i\,m\,\phi}}{\sqrt{2\,\pi}} \qquad l_z = m\,\hbar \qquad m = 0, \pm 1, \pm 2, \dots \qquad (13.106)$$

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.

We insert this form into the L^2 eigenvector-eigenvalue equation. 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

$$-\hbar^2 \left(\frac{1}{\sin\theta} \frac{d}{d\theta} \sin\theta \frac{d}{d\theta} + \frac{\alpha}{\hbar^2} - \frac{m^2}{\sin^2\theta} \right) \Theta_{\alpha,m}(\theta) = 0 \tag{13.107}$$

This is now just a differential equation in θ. Let us change variables to $u = \cos\theta$ and define $P_\alpha^m(u = \cos\theta) = \Theta_{\alpha,m}$. This yields

$$\left\{ \frac{d}{du} \left[(1 - u^2) \frac{d}{du} \right] + \left[\frac{\alpha}{\hbar^2} - \frac{m^2}{1 - u^2} \right] \right\} P_\alpha^m(u) = 0 \qquad -1 \leq u \leq 1 \tag{13.108}$$

If we set $m = 0$, we obtain

$$\left\{ \frac{d}{du} \left[(1 - u^2) \frac{d}{du} \right] + \frac{\alpha}{\hbar^2} \right\} P_\alpha^0(u) = 0 \tag{13.109}$$

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

$$\frac{\alpha}{\hbar^2} = l(l+1) \qquad l = 0, 1, 2, \ldots \tag{13.110}$$

The functions are polynomials containing either even or odd powers of u. They are termed the Legendre polynomials and are denoted by $P_l(u)$. There is a formula, called Rodrigues' Formula, that can be used to generate them:

$$P_l(u) = \frac{1}{2^l \, l!} \frac{d^l}{du^l} \left(u^2 - 1 \right)^l \tag{13.111}$$

The full equation, with $m \neq 0$, is solved by the associated Legendre polynomials, which can be generated from the formula

$$P_l^m(u) = (-1)^m (1 - u^2)^{m/2} \frac{d^m}{du^m} P_l(u) \qquad 0 \leq m \leq l \qquad (13.112)$$

where $m \leq l$ is enforced by the fact that $P_l(u)$ is a lth-order polynomial; hence, any derivatives of order $m + 1$ or larger simply vanish. We may now write $P_l(u)$ as $P_l^0(u)$ based on this formula. For $m < 0$, we see that the θ differential equation is unchanged by the sign of m, so we define

$$P_l^{-m}(u) = P_l^m(u) \qquad 0 \leq m \leq l \qquad (13.113)$$

Combining all of this, we thus write our joint eigenfunctions of L^2 and L_z as

$$Y_l^m(\theta, \phi) = \sqrt{\frac{2\,l+1}{4\,\pi}\frac{(l-m)!}{(l+m)!}}\,P_l^m(u = \cos\theta)\,e^{i\,m\,\phi} \tag{13.114}$$

where the prefactor ensures correct normalization when integrated over all solid angles. These functions are known as the spherical harmonics. The orthonormalization condition (arising from the separate orthonormality of the polar and azimuthal functions) is

$$\int_{-1}^{1} d\cos\theta \int_0^{2\pi} d\phi \left[Y_l^m(\theta, \phi)\right]^* \; Y_{l',m'}(\theta, \phi) = \delta_{ll'}\delta_{mm'} \tag{13.115}$$

The full wavefunction may have any radial dependence as long as its angular dependence is in the form of a spherical harmonic.

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 $\cdot^2$ 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

$$L^2|\alpha, \beta\rangle = \alpha|\alpha, \beta\rangle \qquad L_z|\alpha, \beta\rangle = \beta|\alpha, \beta\rangle \qquad (13.116)$$

We define angular momentum raising and lowering operators:

$$L_\pm = L_x \pm i\,L_y \tag{13.117}$$

They are named this way because they satisfy

$$[L_z, L_\pm] = \pm\hbar\,L_\pm \tag{13.118}$$

so that

$$L_z\left(L_\pm|\alpha,\beta\rangle\right) = \left(\pm\hbar\,L_\pm + L_\pm\,L_z\right)|\alpha,\beta\rangle = \left(\pm\hbar + \beta\right)\left(L_\pm|\alpha,\beta\rangle\right) \tag{13.119}$$

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,

$$[L^2, L_\pm] = 0 \tag{13.120}$$

so we are assured that $|\alpha,\beta\rangle$ and $L_\pm|\alpha,\beta\rangle$ have the same eigenvalue α of L^2.

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

$$L_\pm|\alpha, \beta\rangle = C_\pm(\alpha, \beta)|\alpha, \beta \pm \hbar\rangle \tag{13.121}$$

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:

$$\langle\alpha, \beta|\left(L^2 - L_z^2\right)|\alpha, \beta\rangle = \langle\alpha, \beta|\left(L_x^2 + L_y^2\right)|\alpha, \beta\rangle \tag{13.122}$$

The latter expression is nonnegative because the eigenvalues of L_x^2 and L_y^2 are all nonnegative. So we see $\alpha - \beta^2 \geq 0$, or $\alpha \geq \beta^2$ as desired.

So, we require there to be states $|\alpha, \beta_{max}\rangle$ and $|\alpha, \beta_{min}\rangle$ that satisfy

$$L_+|\alpha, \beta_{max}\rangle = |0\rangle \qquad L_-|\alpha, \beta_{min}\rangle = |0\rangle \tag{13.123}$$

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:

$$L_-L_+|\alpha, \beta_{max}\rangle = L_-|0\rangle \qquad\qquad L_+L_-|\alpha, \beta_{min}\rangle = L_+|0\rangle \tag{13.124}$$

$$\left(L^2 - L_z^2 - \hbar L_z\right)|\alpha, \beta_{max}\rangle = |0\rangle \qquad \left(L^2 - L_z^2 + \hbar L_z\right)|\alpha, \beta_{min}\rangle = |0\rangle \tag{13.125}$$

$$\left(\alpha - \beta_{max}^2 - \hbar\beta_{max}\right)|\alpha, \beta_{max}\rangle = |0\rangle \qquad \left(\alpha - \beta_{min}^2 + \hbar\beta_{min}\right)|\alpha, \beta_{min}\rangle = |0\rangle \tag{13.126}$$

$$\beta_{max}\left(\beta_{max} + \hbar\right) = \alpha \qquad\qquad \beta_{min}\left(\beta_{min} - \hbar\right) = \alpha \tag{13.127}$$

which implies

$$\beta_{min} = -\beta_{max} \tag{13.128}$$

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

$$(L_+)^{k_++1} \, |\alpha, \beta_{min} \rangle \propto |\alpha, \beta_{max} \rangle \qquad (L_-)^{k_-+1} \, |\alpha, \beta_{max} \rangle \propto |\alpha, \beta_{min} \rangle \qquad (13.129)$$

Therefore

$$\beta_{min} + \hbar k_+ = \beta_{max} \qquad \beta_{max} - \hbar k_- = \beta_{min} \qquad (13.130)$$

So we have

$$k_+ = k_- \equiv k \qquad \beta_{max} - \beta_{min} = \hbar k \qquad (13.131)$$

Since $\beta_{min} = -\beta_{max}$, we then have

$$\beta_{max} = k \, \frac{\hbar}{2} \qquad \alpha = \beta_{max} \, (\beta_{max} + \hbar) = \hbar^2 \, \frac{k}{2} \left(\frac{k}{2} + 1 \right) \qquad k = 0, 1, 2, \ldots \quad (13.132)$$

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.)

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.

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.

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 $\beta = m\,\hbar$.

Summary

Let us 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.

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 34:
Operators in the $\{|j, m\rangle\}$ Basis
Relation between the $\{|j, m\rangle\}$ and Coordinate Basis
Representations

Revision Date: 2008/01/23

Operators in the $\{|j, m\rangle\}$ Basis

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

$$J_\pm |\alpha, \beta\rangle = C_\pm(\alpha, \beta)|\alpha, \beta \pm \hbar\rangle \tag{13.133}$$

which we should now rewrite as

$$J_\pm |j, m\rangle = C_\pm(j, m)|j, m + 1\rangle \tag{13.134}$$

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:

$$1 = \langle j, m \pm 1 | j, m \pm 1\rangle = |C_\pm(j, m)|^{-2} |J_\pm |j, m\rangle|^2 \tag{13.135}$$

$$= |C_\pm(j, m)|^{-2} \langle j, m |J_\mp J_\pm|j, m\rangle \tag{13.136}$$

$$= |C_\pm(j, m)|^{-2} \langle j, m| \left(J^2 - J_z^2 \mp \hbar J_z\right)|j, m\rangle \tag{13.137}$$

$$= |C_\pm(j, m)|^{-2} \left[j(j + 1) - m^2 \mp m\right]\hbar^2 \tag{13.138}$$

$$\implies |C_\pm(j, m)|^2 = \hbar^2(j \mp m)(j \pm m + 1) \tag{13.139}$$

We discard the phase freedom and take

$$C_\pm(j, m) = \hbar\sqrt{(j \mp m)(j \pm m + 1)} \tag{13.140}$$

So we are left with

$$J_\pm|j, m\rangle = \hbar\sqrt{(j \mp m)(j \pm m + 1)}|j, m \pm 1\rangle \tag{13.141}$$

We have the expected result that $J_\pm$ annihilates $|j, \pm j\rangle$.

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:

$$\begin{aligned}
\langle j', m'|J_x|j, m\rangle &= \langle j', m'|\left(\frac{J_+ + J_-}{2}\right)|j, m\rangle \tag{13.142} \\
&= \frac{\hbar}{2}\left[\delta_{j,j'}\delta_{m',m+1}\sqrt{(j-m)(j+m+1)}\right. \\
&\qquad \left. + \delta_{j,j'}\delta_{m',m-1}\sqrt{(j+m)(j-m+1)}\right] \tag{13.143}
\end{aligned}$$

Similarly,

$$
\begin{aligned}
\langle j', m' | J_y | j, m \rangle &= \langle j', m' | \left(\frac{J_+ - J_-}{2i} \right) | j, m \rangle \qquad (13.144) \\
&= \frac{\hbar}{2i} \left[\delta_{j,j'} \delta_{m',m+1} \sqrt{(j-m)(j+m+1)} \right. \\
&\qquad \left. - \delta_{j,j'} \delta_{m',m-1} \sqrt{(j+m)(j-m+1)} \right] \qquad (13.145)
\end{aligned}
$$

Operators in the $\{|j, m\rangle\}$ Basis (cont.)

Let us write out the matrix representations so we can get a better feel for how the operators act in the $\{|j, m\rangle\}$ basis. Let's first consider the $j = 1$ subspace for simplicity, which we denote by $V^{(1)}$ and which we will indicate by putting a $^{(1)}$ subscript on the operators. We have

$$
J_z^{(1)} = \hbar \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \qquad [J^2]^{(1)} = 2\,\hbar^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{13.146}
$$

$$
J_x^{(1)} = \frac{\hbar}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad J_y^{(1)} = \frac{i\,\hbar}{\sqrt{2}} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \tag{13.147}
$$

$$
J_+^{(1)} = \hbar\sqrt{2} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad J_-^{(1)} = \hbar\sqrt{2} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{13.148}
$$

Operators in the $\{|j, m\rangle\}$ Basis (cont.)

That is, we have:

▶ $[J^2]^{(1)}$ is diagonal and the $V^{(1)}$ subspace is degenerate for it.

▶ $J_z^{(1)}$ is diagonal and it breaks the degeneracy in the $V^{(1)}$ 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.

▶ $J_x^{(1)}$ and $J_y^{(1)}$ are not diagonal, but we see that they connect states that differ in m by one unit.

▶ $J_+^{(1)}$ and $J_-^{(1)}$ are also 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 $|1, 1\rangle$.

When we consider all possible j values, J^2 and J_z will remain completely diagonal. 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 we saw above when we wrote out their explicit forms: they always have $\delta_{j,j'}$; it is also written out explicitly as matrices in Shankar Section 12.5, p. 328). This 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:

$$V = V^{(0)} \oplus V^{(1/2)} \oplus V^{(1)} \oplus V^{(3/2)} \oplus V^{(2)} \oplus \cdots \tag{13.149}$$

Operators in the $\{|j, m\rangle\}$ Basis (cont.)

Rotation Operators in the $\{|j, m\rangle\}$ Basis

The various J_i operators are the generators of rotations about the i axes, and any rotation can be written in terms of them as $T(\vec{\theta}) = \exp(-(i/\hbar)\,\vec{\theta} \cdot \vec{J})$. 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_i. So we see that rotations cannot change j.

We may write explicit forms for the rotation operators in the $V^{(j)}$ subspaces using the subspace restricted version of $\vec{J}$. In general, we have

$$T^{(j)}(\vec{\theta}) = \exp\left(-\frac{i}{\hbar}\,\vec{\theta} \cdot \vec{J}^{(j)}\right) = \sum_{n=0}^{\infty} \frac{1}{n!}\,\left(-\frac{i}{\hbar}\,\theta\right)^n \left(\hat{\theta} \cdot \vec{J}^{(j)}\right)^n \tag{13.150}$$

This may seem difficult to evaluate, but it turns out not to be because one can show that $\left(\hat{\theta} \cdot \vec{J}^{(j)}\right)^n$ for $n > 2j$ can be written as a linear combination of the first $2j$ powers of $\hat{\theta} \cdot \vec{J}^{(n)}$. This is similar to the way in which the generators of classical rotations, the $\vec{M}$ matrices, satisfied $M_i^{2n} = (-1)^n\,M_i^2$ so that only M_i and M_i^2 were unique and independent.

Therefore, we may write

$$T^{(j)}(\vec{\theta}) = \sum_{n=0}^{2j} f_n(\theta) \left(\widehat{\theta} \cdot \vec{J}^{(j)} \right)^n \tag{13.151}$$

Specific examples are:

$$T^{(0)}(\vec{\theta}) = I^{(0)} \tag{13.152}$$

$$T^{(1/2)}(\vec{\theta}) = \cos\frac{\theta}{2} + 2\,\frac{i}{\hbar}\,\widehat{\theta} \cdot \vec{J}^{(1/2)} \sin\frac{\theta}{2} \tag{13.153}$$

$$T^{(1)}(\theta_i \widehat{i}) = \left[I^{(1)} + (\cos\theta_i - 1)\left(\frac{J_i^{(1)}}{\hbar}\right)^2 \right] - i \sin\theta_i \left(\frac{J_i^{(1)}}{\hbar} \right) \tag{13.154}$$

where i runs over x, y, z and $\widehat{i}$ is the unit vector for he 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.

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 $V^{(j)}$ invariant subspaces: rotations may mix up members of a particular subspace, but they never send a subspace member out of that suspace. "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 $V^{(j)}$ that is smaller than $V^{(j)}$.

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 $V^{(j)}$ could be made smaller, then the subblocks would indicate the subspaces of $V^{(j)}$ that are invariant (closed) under rotations, which we have just concluded can be no smaller than $V^{(j)}$ itself.

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 really 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 Eigensubspaces of Rotationally Invariant Hamiltonians

Let us consider Hamiltonians that satisfy $[H, \vec{J}] = 0$. We term these rotationally invariant because the above condition implies that a rotation transformation about any axis will by 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 eigensubspaces of a rotationally invariant H. That is, all elements of a subspace $V^{(j)}$ 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 $J_\pm^k |j, m\rangle = |0\rangle$). One can reach any of the $\{|j, m\rangle\}$ in $V^{(j)}$ using enough powers of $J_\pm$, so all the $\{|j, m\rangle\}$ in $V^{(j)}$ must also be eigenstates of H of energy E. Since the $\{|j, m\rangle\}$ span the $V^{(j)}$ subspace, $V^{(j)}$ is thus an eigensubspace of H with eigenvalue E.

Relation between $|j, m\rangle$ Basis and Position Basis Eigenstates

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 $\{|l, m\rangle\}$. 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 is specific to orbital angular momentum. Such a basis simply does not exist for spin.

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 $|l, l\rangle$ and the lowering operator annihilates $|l, -l\rangle$. We shall see that we need to us both relations, so let's begin with both:

$$L_\pm |l, \pm l\rangle = |0\rangle \tag{13.155}$$

Let's project onto the position basis:

$$\langle r, \theta, \phi | (L_x \pm i\, L_y) |l, \pm l\rangle = \langle r, \theta, \phi |0\rangle = 0 \tag{13.156}$$

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

$$\pm \hbar e^{i\phi} \left(\frac{\partial}{\partial \theta} \pm i \cot\theta \, \frac{\partial}{\partial \phi} \right) \psi_l^{\pm l}(r, \theta, \phi) = 0 \tag{13.157}$$

Relation between $|j, m\rangle$ Basis and Position Basis Eigenstates (cont.)

Since we know $\psi_l^{\pm l}(r, \theta, \phi)$ must be an eigenstate of L_z with eigenvalue $\pm l\,\hbar$, we know the solution must be of the form

$$\psi_l^{\pm l}(r, \theta, \phi) = U_l^{\pm l}(r, \theta)\, e^{\pm i\, l\, \phi} \tag{13.158}$$

Inserting this, we obtain

$$\left(\frac{d}{d\theta} - l \cot\theta\right) U_l^{\pm l}(r, \theta) = 0 \tag{13.159}$$

where we canceled out the nonvanishing $e^{\pm i\phi}$ and $e^{\pm i\, l\, \phi}$ factors. This is integrable:

$$\frac{dU_l^{\pm l}}{U_l^{\pm l}} = l\, \frac{d(\sin\theta)}{\sin\theta} \tag{13.160}$$

$$U_l^{\pm l} = R(r)\, (\sin\theta)^l \tag{13.161}$$

The angular part of $\psi_l^{\pm l}$ ought to be the spherical harmonic $Y_l^{\pm l}$, after we correct for normalization and follow the same sign convention as we used before. It is:

$$Y_l^{\pm l}(\theta, \phi) = (-1)^l \sqrt{\frac{(2\,l+1)!}{4\,\pi}} \, \frac{1}{2^l \, l!} \, (\sin\theta)^l \, e^{\pm i\,l\,\phi} \qquad (13.162)$$

That is, $Y_l^{\pm l}$ has the same θ and ϕ dependences as the solutions we found to the annihilation conditions.

We of course obtain the position-basis representations by using the lowering and raising operators on the $|l, \pm l\rangle$ states:

$$\psi_l^{\pm m}(r, \theta, \phi) \tag{13.163}$$
$$= \langle r, \theta, \phi \,|l, \pm m\rangle$$
$$= \langle r, \theta, \phi \,|L_{\mp}^{(l-m)}|l, \pm l\rangle \tag{13.164}$$
$$= \left[\mp \hbar e^{\mp i\phi} \left(\frac{\partial}{\partial \theta} \mp i \cot\theta \frac{\partial}{\partial \phi}\right)\right]^{(l-m)} R(r)\, Y_l^{\pm l}(\theta, \phi) \tag{13.165}$$
$$= R(r)\,(-1)^l (\pm 1)^m \sqrt{\frac{2\,l+1}{4\,\pi} \frac{(l+m)!}{(l-m)!}}\; e^{\pm i\,m\,\phi} (\sin\theta)^{-m} \left(\frac{d}{d(\cos\theta)}\right)^{l-m} (\sin\theta)^{2\,l} \tag{13.166}$$
$$= R(r)\, Y_l^{\pm m} \tag{13.167}$$

where we recognize Rodrigues' formula in the penultimate step. We recover the spherical harmonics completely.

We note that we used the annihilation conditions on both $|l, \pm l\rangle$ 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 35:
Rotationally Invariant Problems in Three Dimensions

Revision Date: 2008/03/02

Rotationally Invariant Problems in Three Dimensions

Requirements for $[H, L_z] = 0$ and $[H, L^2] = 0$

With the above solutions to the eigenvector-eigenvalue problem of L^2 and L_z in hand, we may simplify the Schrödinger Equation for problems for which $[H, L_z] = 0$ and $[H, L^2] = 0$ in three dimensions.

Let's first determine the condition for $[H, L_z] = 0$ and $[H, L^2] = 0$ in three dimensions. We have already shown that $[P_x^2 + P_y^2, L_z] = 0$. We are assured $[P_z, L_z] = 0$ by the commutation relations we gave earlier. One can also see this by recognizing that $L_z = X P_y - Y P_x$; all these operators commute with P_z. So we are assured that $[T, L_z] = 0$, that the kinetic energy term commutes with L_z. Similar logic would show that $[T, L_x] = 0$ and $[T, L_y] = 0$: there is nothing special about L_z. Since $[T, L_i] = 0$ for all i, and $L^2 = L_x^2 + L_y^2 + L_z^2$, we are thereby certain that $[T, L^2] = 0$. So no condition need be placed on the kinetic energy to ensure it commutes with both L_z and L^2.

For the potential, $[V, L_z] = 0$ implies that V may not depend on ϕ, only $V = V(r, \theta)$ is allowed. So we are now left with the requirement $[V, L^2] = 0$. Once can see, however, that this does not imply that V has no θ dependence. Let's try integrating the requirement $(L^2 V) = 0$. It is easier to do in terms of $u = \cos\theta$, so let's use the form we found above in connection with the derivation of the spherical harmonics:

$$L^2 V(r, u = \cos\theta) = 0 \tag{13.168}$$

$$\frac{\partial}{\partial u}\left[\left(1 - u^2\right)\frac{\partial}{\partial u}\right] V(r, u) = 0 \tag{13.169}$$

We may integrate over u to obtain

$$\left[\left(1 - u^2\right) \frac{\partial}{\partial u}\right] V(r, u) = A(r) \tag{13.170}$$

where $A(r)$ is a purely r-dependent function (the "constant" of integration may depend on r). We may rewrite as

$$\frac{\partial}{\partial u} V(r, u) = \frac{A(r)}{1 - u^2} \tag{13.171}$$

$$V(r, u) = B(r) + \frac{C(r)}{2} \log \frac{1 + u}{1 - u} \tag{13.172}$$

$$V(r, \theta) = B(r) + \frac{C(r)}{2} \log \frac{1 + \cos \theta}{1 - \cos \theta} \tag{13.173}$$

So a very specific type of θ dependence is allowed, and it results in a θ dependence that blows up at $\theta = 0$ and $\theta = \pi$.

Rotationally Invariant Problems

While the above θ dependence is allowed by $[H, L_z] = 0$ and $[H, L^2] = 0$, we should not retain this level of generality because these two commutation conditions are not what we mean by rotational invariance in three dimensions. Rotational invariance in 3d 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.

In order for $[H, \vec{L}] = 0$, we need $[V, \vec{L}] = 0$. As we have already shown, $[H, L_z] = 0$ implies $V = V(r, \theta)$. With this in hand one then easily sees $[H, L_x] = 0$ or $[H, L_y] = 0$ implies $V = V(r)$ because these conditions reduce to requirements that $\frac{\partial V}{\partial \theta} = 0$ since the ϕ dependence has already been knocked out.

So, we see that rotational invariance implies the potential is spherically symmetric, $V = V(r)$. We shall consider this case here.

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

$$
\left\{-\frac{\hbar^2}{2\,\mu}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right] + V(r)\right\}\psi_E(r,\theta,\phi)
$$
$$
= E\,\psi_E(r,\theta,\phi) \tag{13.174}
$$

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

$$
\psi_E(r,\theta,\phi) = R_{E,l}(r)Y_l^m(\theta,\phi) \tag{13.175}
$$

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

$$
\left\{-\frac{\hbar^2}{2\,\mu}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{l(l+1)}{r^2}\right] + V(r)\right\}R_{E,l}(r) = E\,R_{E,l}(r) \tag{13.176}
$$

This is the radial equation for a spherically symmetric potential in three dimensions.

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

$$R_{E,l}(r) = \frac{U_{E,l}(r)}{r} \tag{13.177}$$

because the radial equation then simplifies to the reduced radial equation

$$D_l(r)\, U_{E,l}(r) \equiv \left[-\frac{\hbar^2}{2\,\mu}\, \frac{d^2}{dr^2} + V(r) + \frac{l\,(l+1)\,\hbar^2}{2\,\mu\,r^2} \right] U_{E,l}(r) = E\, U_{E,l}(r) \tag{13.178}$$

This looks like the Schrödinger Equation in one dimension for a potential

$$V_{eff}(r) = V(r) + \frac{l\,(l+1)\,\hbar^2}{2\,\mu\,r^2} \tag{13.179}$$

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 1-D Schrödinger Equation defined over an interval, and we have the Hermitian operator $D_l(r)$, we must check that the standard Hermiticity boundary condition is satisfied. That is, we require

$$\int_0^\infty dr \, U_1^* \left(D_l U_2\right) = \left[\int_0^\infty dr \, U_2^* \left(D_l U_1\right)\right]^* \tag{13.180}$$

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

$$\int_0^\infty dr \, U_1^* \frac{d^2}{dr^2} \, U_2 = \int_0^\infty dr \, U_2 \frac{d^2}{dr^2} \, U_1^* \tag{13.181}$$

Let's manipulate the RHS:

$$RHS = \int_0^\infty dr \, U_2 \frac{d^2}{dr^2} U_1^* = U_2 \frac{d}{dr} U_1^* \Big|_0^\infty - \int_0^\infty dr \frac{dU_2}{dr} \frac{dU_1^*}{dr} \tag{13.182}$$

$$= U_2 \frac{dU_1^*}{dr} \Big|_0^\infty - \frac{dU_2}{dr} U_1^* \Big|_0^\infty + \int_0^\infty dr \frac{d^2 U_2}{dr^2} U_1^* \tag{13.183}$$

So, to obtain $LHS = RHS$, we require

$$\left[U_2 \frac{dU_1^*}{dr} - \frac{dU_2}{dr} U_1^* \right] \Big|_0^\infty = 0 \tag{13.184}$$

$$\frac{d}{dr} \left[U_2 \, U_1^* \right] \Big|_0^\infty = 0 \tag{13.185}$$

To evaluate the ∞ limit of the above, we need to know the behavior of $U_{E,l}(r)$ 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

$$\int_0^\infty dr \, r^2 \, |R_{E,l}(r)|^2 = \int_0^\infty dr \, |U_{E,l}(r)|^2 \tag{13.186}$$

In order for this to equal unity, we need $U_{E,l}(r) \to 0$ as $r \to \infty$. To make it normalizable to a delta function, we require $U_{E,l}(r) \to e^{ikr}$ 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

$$\frac{d}{dr} \left[U_2 \, U_1^* \right] \bigg|_0 = 0 \tag{13.187}$$

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

$$U(r) \xrightarrow{r \to 0} c \tag{13.188}$$

with c a constant.

Form of Solutions for $r \to 0$

We have established that, in order for the $D_l(r)$ operator to be Hermitian, we require $U_{E,l}(r) \to c$ as $r \to 0$ and $U_{E,l}(r) \to 0$ or e^{ikr} as $r \to \infty$. Now let us check whether these requirements are consistent with the Schrödinger Equation.

Let's first check $r \to 0$. It is insufficient to check the one-dimensional equation for $U_{E,l}(r)$ because the relation $R_{E,l}(r) = U_{E,l}(r)/r$ breaks down at $r \to 0$ unless $U_{E,l}(r) \to 0$ as fast as r. So we need to check the full three-dimensional Schrödinger Equation. That is

$$-\frac{\hbar^2}{2\,\mu}\,\nabla^2 \psi_{E,l}(r,\theta\,\phi) + V(r)\,\psi_{E,l}(r,\theta,\phi) = E\,\psi_{E,l}(r,\theta,\phi) \tag{13.189}$$

which, based on our asymptotic form for $r \to 0$, reduces to

$$\left[-\frac{\hbar^2}{2\,\mu}\,\nabla^2 + V(r) + \frac{\hbar^2 l\,(l+1)}{2\,\mu\,r^2} \right] \frac{1}{r} = E\,\frac{1}{r} \tag{13.190}$$

This is problematic in two ways. First, for $r \neq 0$, $\nabla^2(1/r) = 0$ for $r \neq 0$, so we obtain

$$V(r) \xrightarrow{r \to 0} E - \frac{\hbar^2 l(l+1)}{2\mu r^2} \tag{13.191}$$

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:

$$\int_{\mathcal{V}(r)} d\Omega \, (r')^2 \, dr' \, \nabla^2 \frac{1}{r'} = \int_{\mathcal{S}(r)} d\Omega \, r^2 \, \hat{r} \cdot \vec{\nabla} \frac{1}{r'} \tag{13.192}$$

$$\int_{\mathcal{V}(r)} d\Omega \, (r')^2 \, dr' \, \nabla^2 \frac{1}{r'} = 4\pi r^2 \frac{d}{dr} \frac{1}{r} \tag{13.193}$$

$$\int_{\mathcal{V}(r)} d\Omega \, (r')^2 \, dr' \, \nabla^2 \frac{1}{r'} = -4\pi \tag{13.194}$$

Generic Properties of Solutions of the Radial Equation (cont.)

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

$$\nabla^2 \frac{1}{r} = 4\pi\,\delta(\vec{r}) \equiv 4\pi\,\delta(x)\,\delta(y)\,\delta(z) \tag{13.195}$$

Plugging this back into the Schrödinger Equation, we see the Schrödinger Equation can only be satisfied if $V(r) \to \delta(\vec{r})$ 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 $r \to 0$, $U_{E,l}(r) \to 0$. How quickly $U_{E,l}(r)$ must vanish will be studied next.

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 $l \neq 0$, 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

$$\frac{d^2}{dr^2} U_l(r) = \frac{l(l+1)}{r^2} U_l(r) \tag{13.196}$$

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 $U_l(r) = r^\gamma$ implies

$$\gamma(\gamma-1) = l(l+1) \qquad \Longrightarrow \qquad \gamma = l+1 \quad \text{or} \quad \gamma = -l \tag{13.197}$$

also known as the *regular* and *irregular* solutions because of their behavior near the origin. The latter one fails our condition $U_{E,l}(r) \to 0$ for $r \to 0$, so we keep only the regular solution. Since $l \geq 1$, we are assured that $R_{E,l}(r) = U_{E,l}(r)/r \to 0$ 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 $l = 0$, 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. We will show later that the above form $U_l(r) = r^{l+1}$ is also valid for $l = 0$ for the Coulomb potential. This results in $R_{E,l}(r) \to r/r = 1$ as $r \to 0$, so the absence of a centrifugal barrier allows the particle to be found at the origin.

Generic Properties of Solutions of the Radial Equation (cont.)

Form of Solutions for $r \rightarrow \infty$

Let's consider potentials for which $V(r) \rightarrow 0$ as $r \rightarrow \infty$ 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 one-dimensional Schrödinger Equation reduces to

$$\frac{d^2}{dr^2} U_E(r) = -\frac{2\,\mu\,E}{\hbar^2}\, U_E(r) \tag{13.198}$$

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.

For $E > 0$, the exponential has an imaginary argument and our solutions are of the form

$$U_E(r) = A\,e^{ikr} + B\,e^{-ikr} \qquad k = \frac{1}{\hbar}\sqrt{2\,\mu\,E} \tag{13.199}$$

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^{l+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.

That is, we start with the standard WKB form

$$U_E(r) = e^{\pm \frac{i}{\hbar} \phi(r)} \qquad \phi(r) = \int_{r_0}^r dr \sqrt{2\,\mu\,(E - V(r))} \tag{13.200}$$

(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:

$$\phi(r) = \int_{r_0}^r dr' \sqrt{2\,\mu\,E} \left(1 - \frac{1}{2}\,\frac{V(r')}{E} \right) = \hbar k \int_{r_0}^r dr' - \frac{\hbar k}{2\,E} \int_{r_0}^r dr'\, V(r') \tag{13.201}$$

$$= \hbar k\,(r - r_0) - \frac{\mu}{\hbar k} \int_{r_0}^r dr'\, V(r') \tag{13.202}$$

So we have

$$U_E(r) = f(r_0)\, e^{\pm i k r}\, \exp\left(\mp \frac{i}{\hbar}\,\frac{\mu}{\hbar k} \int_{r_0}^r dr'\, V(r') \right) \tag{13.203}$$

where $f(r_0)$ is a normalization factor that depends on the choice of lower limit of the integral.

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

$$\int_{r_0}^{\infty} dr' \, V(r') = c(r_0) \tag{13.204}$$

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

$$r \, V(r) \overset{r \to \infty}{\longrightarrow} 0 \tag{13.205}$$

(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

$$U_E(r) \to e^{-\kappa r} \qquad \kappa = \frac{1}{\hbar}\sqrt{2\,\mu\,|E|} \qquad (13.206)$$

We may repeat the WKB procedure above to find

$$U_E(r) = f(r_0)\, e^{-\kappa r}\, \exp\left(-\frac{1}{\hbar}\frac{\mu}{\hbar\kappa}\int_{r_0}^{r} dr'\, V(r')\right) \qquad (13.207)$$

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:

$$\int_{r_0}^{r} dr'\, V(r') = -e^2 \int_{r_0}^{r} dr'\, \frac{1}{r'} = -e^2\,(\ln r - \ln r_0) \qquad (13.208)$$

$$\implies \qquad U_E(r) = g(r_0)\, r^{\,\mu\, e^2/\hbar^2\kappa}\, e^{-\kappa r} \qquad (13.209)$$

(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.)

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 $U \xrightarrow{r \to 0} 0$ are nondegenerate in the same way that we showed that bound states of the one-dimensional Schrödinger Equation are nondegenerate.

Hence, we are assured that there is an orthonormalization condition for bound states

$$\int_0^\infty dr \, U_{E,l}(r) \, U_{E',l}(r) = \delta_{E,E'} \tag{13.210}$$

or, using $\psi_{E,l,m}(r) = R_{E,l}(r) \, Y_l^m(\theta, \phi)$, the orthonormalization condition for the full 3-dimension bound eigenstates is

$$\int_0^\infty r^2 \, dr \int_0^\pi \sin\theta \, d\theta \int_0^{2\pi} d\phi \, \psi_{E,l,m}^*(r,\theta,\phi) \, \psi_{E',l',m'}(r,\theta,\phi) = \delta_{E,E'} \delta_{l,l'} \delta_{m,m'} \tag{13.211}$$

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.

Solutions for Specific Rotationally Invariant Potentials

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

$$\left[\frac{d^2}{dr^2} + k^2 - \frac{l(l+1)}{r^2} \right] U_{E,l}(r) = 0 \qquad k = \frac{1}{\hbar}\sqrt{2\,\mu\,E} \qquad (13.212)$$

One can solve the equation easily for $l = 0$, 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 $l = 0$ case, directly, and use the raising operators to obtain the $l \neq 0$ solutions.

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

$$j_0(\rho) = \frac{\sin \rho}{\rho} \qquad\qquad \eta_0(\rho) = -\frac{\cos \rho}{\rho} \qquad\qquad (13.213)$$

$$j_1(\rho) = \frac{\sin \rho}{\rho^2} - \frac{\cos \rho}{\rho} \qquad\qquad \eta_1(\rho) = -\frac{\cos \rho}{\rho^2} - \frac{\sin \rho}{\rho} \qquad\qquad (13.214)$$

These functions have asymptotic forms

$$j_l(\rho) \xrightarrow{\rho \to 0} \frac{\rho^l}{(2\,l+1)!!} \qquad\qquad \eta_l(\rho) \xrightarrow{\rho \to 0} -\frac{(2\,l-1)!!}{\rho^{l+1}} \qquad\qquad (13.215)$$

$$j_l(\rho) \xrightarrow{\rho \to \infty} \frac{1}{\rho} \sin\left(\rho - l\,\frac{\pi}{2}\right) \qquad \eta_l(\rho) \xrightarrow{\rho \to \infty} -\frac{1}{\rho} \cos\left(\rho - l\,\frac{\pi}{2}\right) \qquad\qquad (13.216)$$

Only the spherical Bessel functions are regular as $\rho \to 0$, 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

$$\psi_{E,l,m}(r, \theta, \phi) = j_l(k\,r)\, Y_l^m(\theta, \phi) \qquad k = \frac{1}{\hbar}\sqrt{2\,\mu\,E} \qquad (13.217)$$

with orthonormalization condition

$$\int_0^\infty r^2 dr \int_0^\pi \sin\theta\, d\theta \int_0^{2\pi} d\phi\; \psi_{E,l,m}^*(r, \theta, \phi)\, \psi_{E',l',m'}(r, \theta, \phi)$$
$$= \frac{2}{\pi\,k^2}\, \delta(k - k')\delta_{l,l'}\delta_{m,m'} \qquad (13.218)$$

Section 14
Spin Angular Momentum

Lecture 36:
Tensors in Classical and Quantum Mechanics

Revision Date: 2008/01/28

Review of 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, so we may define a sensible extension in quantum mechanics.

The material discussed here is cribbed directly from Section 5.1.4 of the Ph106ab Lecture Notes, but we adapt it slightly to the background material provided so far in this course and drop some examples that are not relevant here.

Tensors — Definitions

We have discussed passive coordinate transformations of quantum mechanical states. The idea there was that we have untransformed and transformed sets of coordinate axes, with their associated position bases $\{|q\rangle\}$ and $\{|q'\rangle\}$, with some relationship between the two bases, and that we developed rules for how the representations of some state $|\psi\rangle$ in terms of the two bases are related.

Here, we consider a physical 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. The transformation rules between coordinate representations are not the same as for quantum mechanical states, but the concept is similar.

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, 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 classical rotation matrices we discussed in Sections 13.2 and 13.4.

You are certainly familiar with two kinds of tensors. The first, called a scalar, is also known as a rank 0 tensor. 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 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 $\{v_i\}$. Its representation in a different frame F', which we shall denote by $\underline{\vec{v}}'$ or $\{v_i'\}$, is related to that in F by

$$\underline{\vec{v}}' = \mathbf{R}_{P,\vec{\theta}}\, \underline{\vec{v}} \tag{14.1}$$

$$v_i' = \sum_i \left(\mathbf{R}_{P,\vec{\theta}}\right)_{ij} v_i \tag{14.2}$$

where $\mathbf{R}_{P,\vec{\theta}}$ is the *passive* rotation matrix that relates F and F'. Recall that, by convention, $R_{P,\vec{\theta}} = R_{A,-\vec{\theta}} \equiv R_{-\vec{\theta}}$.

Let us be more specific about how $\mathbf{R}_{-\vec{\theta}}$ is defined because there can be significant confusion about the sign of the rotation angle involved. Consider the unit vectors of the F' system, which we will denote by $\{\vec{e}_i'\}$, $i = 1, 2, 3$. (This time, the prime refers to the frame for which these are unit vectors, not the frame in which we are showing the representation — hence the lack of underline.) Their coordinate representations in the F' system are, obviously,

$$\underline{\left(\vec{e}_1'\right)'} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \underline{\left(\vec{e}_2'\right)'} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \underline{\left(\vec{e}_3'\right)'} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad (14.3)$$

where the underline indicates "coordinate representation" and the outermost prime tells us this is the coordinate representation in the F' frame. The representations of the $\{\vec{e}_i'\}$ in F are denoted by $\underline{\left(\vec{e}_i'\right)}$ and will, obviously, not be so simple. In order to specify how F' is oriented relative to F, these must have been provided. The $\{\vec{e}_i'\}$ are vectors, too, so for each one it holds that

$$\underline{\left(\vec{e}_i'\right)'} = \mathbf{R}_{-\vec{\theta}}\underline{\left(\vec{e}_i'\right)} \qquad (14.4)$$

More explicitly,

$$\left[\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right] = \mathbf{R}_{-\vec{\theta}}\,\underline{(\vec{e}_1')} \qquad \left[\begin{array}{c} 0 \\ 1 \\ 0 \end{array}\right] = \mathbf{R}_{-\vec{\theta}}\,\underline{(\vec{e}_2')} \qquad \left[\begin{array}{c} 0 \\ 0 \\ 1 \end{array}\right] = \mathbf{R}_{-\vec{\theta}}\,\underline{(\vec{e}_3')} \quad (14.5)$$

Conversely, we denote the unit vectors of F by $\{\vec{e}_i\}$; their representations in F are

$$\underline{(\vec{e}_1)} = \left[\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right] \qquad \underline{(\vec{e}_2)} = \left[\begin{array}{c} 0 \\ 1 \\ 0 \end{array}\right] \qquad \underline{(\vec{e}_3)} = \left[\begin{array}{c} 0 \\ 0 \\ 1 \end{array}\right] \qquad (14.6)$$

The representations in the two frames are related by

$$\underline{(\vec{e}_i)}' = \mathbf{R}_{-\vec{\theta}}\underline{(\vec{e}_i)} \qquad (14.7)$$

$$\underline{(\vec{e}_1)}' = \mathbf{R}_{-\vec{\theta}}\left[\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right] \qquad \underline{(\vec{e}_2)}' = \mathbf{R}_{-\vec{\theta}}\left[\begin{array}{c} 0 \\ 1 \\ 0 \end{array}\right] \qquad \underline{(\vec{e}_3)}' = \mathbf{R}_{-\vec{\theta}}\left[\begin{array}{c} 0 \\ 0 \\ 1 \end{array}\right] \qquad (14.8)$$

Either of these relations is sufficient to specify $\mathbf{R}_{-\vec{\theta}}$, and certainly to define the sign of the rotation angle.

A rank n 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 tensor $\mathcal{T}$ is an object that has coordinate representation $\underline{\mathcal{T}}$ with N^n components $T_{i_1 \cdots i_n}$ (where N is the dimensionality of the physical space, $N = 3$ for what we are considering) with transformation properties

$$T'_{i_1 \cdots i_n} = \sum_{j_1, j_2, \cdots, j_n} R_{i_1 j_1} \cdots R_{i_n j_n} T_{j_1 \cdots j_n}$$

(We drop the $_{-\vec{\theta}}$ for brevity.) We see why a scalar is a rank 0 tensor and a vector is a rank 1 tensor. 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:

$$T'_{ij} = \sum_{j,l} R_{ik} R_{jl} T_{kl} = \sum_{j,l} R_{ik} T_{kl} R^T_{lj} \tag{14.9}$$

$$\underline{\mathcal{T}}' = \mathbf{R} \, \underline{\mathcal{T}} \, \mathbf{R}^T = \mathbf{R} \, \underline{\mathcal{T}} \, \mathbf{R}^{-1} \tag{14.10}$$

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. The last expression is the similarity transformation of the $N \times N$ matrix $\underline{\mathcal{T}}$ by the orthogonal matrix $\mathbf{R}$.

Tensors — Examples

▶ The norm of a vector is a scalar:

$$\underline{\vec{v}}\,' \cdot \underline{\vec{v}}\,' = \sum_i v_i{'}\, v_i{'} = \sum_{i,j,k} R_{ij}\, v_j\, R_{ik}\, v_k = \sum_{j,k} \delta_{jk}\, v_j\, v_k = \sum_j v_j\, v_j = \underline{\vec{v}} \cdot \underline{\vec{v}} \quad (14.11)$$

where we have used $\underline{\vec{v}} = \mathbf{R}^T\, \underline{\vec{v}}\,'$ and the orthonormality property $R_{ij}\, R_{ik} = \delta_{jk}$ (*i.e.*, $\mathbf{R}^T = \mathbf{R}^{-1}$). We see that the value of the norm of a vector is the same in th 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 tensor is the outer product of two vectors:

$$T_{ij} = a_i\, b_j \qquad \text{or} \qquad \mathcal{T} = \vec{a}\, \vec{b}^T \qquad (14.12)$$

Since each vector transforms as a rank 1 tensor, it is obvious that the above product transforms as a rank 2 tensor.

▶ More generally, if we take a rank m tensor with coordinate representation components $U_{i_1 \cdots i_m}$ and a rank n tensor with coordinate representation components $V_{j_1 \cdots j_n}$ 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 $m + n - 2p$ tensor. Proving it is clearly a tedious exercise in index arithmetic relying on the rotation matrix orthogonality relation $R_{ki}\, R_{kj} = \delta_{ij}$ and its transpose relation $R_{ik}\, R_{jk} = \delta_{ij}$. Taking $p = 0$ as a special case gives us the simple outer product of the two tensors, which gives the previous example when both tensors are rank 1.

▶ The identity matrix is a rank 2 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:

$$\mathsf{I}'_{ij} = R_{ik}\, R_{jl}\, \mathsf{I}_{kl} = R_{ik}\, R_{jl}\, \delta_{kl} = R_{ik}\, R_{jk} = \delta_{ij} \tag{14.13}$$

(We used the "transposed" orthonormality condition $R_{ik}\, R_{jk} = \delta_{ij}$.) So, we see that the identity matrix has representation δ_{ij} in any frame and that the representations in different frames are related by the appropriate transformation relations.

▶ We can demonstrate that the ϵ_{ijk} Levi-Civita symbol is an isotropic rank 3 tensor. Let's explicitly calculate the effect of the transformation rule on it:

$$\epsilon'_{ijk} = R_{il}\, R_{jm}\, R_{kn}\, \epsilon_{lmn}$$

We may evaluate the above by recognizing that the "transposed" orthonormality condition on **R** implies that the rows of **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 $\vec{\mathbf{R}}_i^r$, where $\left(\vec{\mathbf{R}}_i^r\right)_j = R_{ij}$. (The ^{r} superscript indicates we are treating the rows, rather than the columns, of **R** as vectors.) With this notation, the above product looks like $\vec{\mathbf{R}}_i^r \cdot \left(\vec{\mathbf{R}}_j^r \times \vec{\mathbf{R}}_k^r\right)$. In $N = 3$ dimensions, the expression will only be nonvanishing when the triplet ijk 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_{ijk}, so we have

$$\epsilon'_{ijk} = \epsilon_{ijk} \tag{14.14}$$

So the Levi-Civita symbol is an isotropic rank 3 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 tensor $\mathcal{M}$ with coordinate representation components $M_{ijk} = \left(\vec{\mathbf{M}}_i\right)_{jk} = -\epsilon_{ijk}$, $\mathcal{M}$ is clearly an isotropic rank 3 tensor. For this particularly interesting case, we will take the symbol $\vec{\mathbf{M}}$ to stand for the rank 3 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 $\vec{\theta}$, the quantity $\vec{\theta} \cdot \vec{\mathbf{M}}$ has in frames F and F' coordinate representations $\underline{\vec{\theta}} \cdot \underline{\vec{\mathbf{M}}} = \underline{\vec{\theta}} \cdot \vec{\mathbf{M}}$ and $\left(\vec{\theta} \cdot \vec{\mathbf{M}}\right)' = \underline{\vec{\theta}}' \cdot \underline{\vec{\mathbf{M}}}' = \underline{\vec{\theta}}' \cdot \vec{\mathbf{M}}$, 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.

Tensors in Quantum Mechanics

Tensor Operators

The first way in which tensors are applied in quantum mechanics is simply a direct mapping of tensor physical quantities in classical mechanics to tensor operators in quantum mechanics. Examples include the vector position operator, $\vec{R} = \hat{x}\,X + \hat{y}\,Y + \hat{z}\,Z = \sum_i \vec{e}_i R_i$, the vector linear momentum operator $\vec{P} = \sum_i \vec{e}_i P_i$, and the vector orbital angular momentum operator $\vec{L} = \vec{R} \times \vec{P}$. We have seen hints of the vectorial nature of these operators in the way they behave under coordinate transformations; see Section 13.2, where we calculated the transformation properties of the components of $\vec{R}$ and $\vec{P}$ under a two-dimensional rotation. Here, we will formalize this behavior.

Tensors in Quantum Mechanics (cont.)

Let's write these transformation laws out explicitly to see how they relate to the classical laws. Consider a rotation transformation

$$T(\vec{\theta}) = T\left(\sum_i \vec{e}_i \theta_i\right) = \exp\left(-\frac{i}{\hbar} \vec{\theta} \cdot \vec{L}\right) \tag{14.15}$$

where the transformed (F') axes are obtained from the untransformed (F) axes by a rotation CCW by θ about $\widehat{\theta}$. We saw that the operators transform via the the relation

$$O' = T(\vec{\theta}) \, O \, T^{\dagger}(\vec{\theta}) \tag{14.16}$$

For example,

$$X' = \exp\left(-\frac{i}{\hbar} \vec{\theta} \cdot \vec{L}\right) X \exp\left(\frac{i}{\hbar} \vec{\theta} \cdot \vec{L}\right) \tag{14.17}$$

We showed for the case of a rotation about the z axis that this resulted in

$$X' = X \cos\theta + Y \sin\theta \qquad Y' = -X \sin\theta + Y \cos\theta \qquad Z' = Z \tag{14.18}$$

Tensors in Quantum Mechanics (cont.)

Let us define two operators

$$\vec{R} = \hat{x}\,X + \hat{y}\,Y + \hat{z}\,Z = \vec{e}_1 R_1 + \vec{e}_2 R_2 + \vec{e}_3 R_3 \tag{14.19}$$

$$\vec{R}\,' = \hat{x}'\,X' + \hat{y}'\,Y' + \hat{z}'\,Z' = \vec{e}_1'R_1' + \vec{e}_2'R_2' + \vec{e}_3'R_3' \tag{14.20}$$

We define these operators in this way because we of course expect we will be able to call these vector operators. But we need to show they transform like vectors. In fact, we will show that $\vec{R}$ and $\vec{R}\,'$, while being distinct operators as so far defined, are tied to each other by the classical vector transformation law. They are *different coordinate representations* of the *same vector operator*.

What do we mean by the coordinate representation of a vector operator? Simply put, just treat the operators as numbers and write out the coordinate representation in the usual way. Since we know the coordinate representations of the $\{\vec{e}_i\}$ and $\{\vec{e}_i'\}$ from our classical discussion of tensors, we may write

$$(\underline{\vec{R}}) = \underline{(\vec{e}_1)}X + \underline{(\vec{e}_2)}Y + \underline{(\vec{e}_3)}Z \tag{14.21}$$

$$= \left[\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right]X + \left[\begin{array}{c} 0 \\ 1 \\ 0 \end{array}\right]Y + \left[\begin{array}{c} 0 \\ 0 \\ 1 \end{array}\right]Z = \left[\begin{array}{c} X \\ Y \\ Z \end{array}\right] \tag{14.22}$$

Similarly,

$$\left(\underline{\vec{R}\,'}\right)' = \left(\underline{\vec{e}_1'}\right)' X' + \left(\underline{\vec{e}_2'}\right)' Y' + \left(\underline{\vec{e}_3'}\right)' Z' \tag{14.23}$$

$$= \left[\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right] X' + \left[\begin{array}{c} 0 \\ 1 \\ 0 \end{array}\right] Y' + \left[\begin{array}{c} 0 \\ 0 \\ 1 \end{array}\right] Z' = \left[\begin{array}{c} X' \\ Y' \\ Z' \end{array}\right] \tag{14.24}$$

Now, the tricky part, and the one that connects our quantum and classical transformation laws, is to write the other possible coordinate representations:

$$\left(\underline{\vec{R}}\right)' = \left[\mathbf{R}_{-\vec{\theta}}(\underline{\vec{e}_1})\right] X + \left[\mathbf{R}_{-\vec{\theta}}(\underline{\vec{e}_2})\right] Y + \left[\mathbf{R}_{-\vec{\theta}}(\underline{\vec{e}_3})\right] Z \tag{14.25}$$

$$= \left(\mathbf{R}_{-\vec{\theta}}\left[\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right]\right) X + \left(\mathbf{R}_{-\vec{\theta}}\left[\begin{array}{c} 0 \\ 1 \\ 0 \end{array}\right]\right) Y + \left(\mathbf{R}_{-\vec{\theta}}\left[\begin{array}{c} 0 \\ 0 \\ 1 \end{array}\right]\right) Z \tag{14.26}$$

$$\left(\underline{\vec{R}\,'}\right) = \left[\mathbf{R}_{\vec{\theta}}(\underline{\vec{e}_1'})'\right] X' + \left[\mathbf{R}_{\vec{\theta}}(\underline{\vec{e}_2'})'\right] Y' + \left[\mathbf{R}_{\vec{\theta}}(\underline{\vec{e}_3'})'\right] Z' \tag{14.27}$$

$$= \left(\mathbf{R}_{\vec{\theta}}\left[\begin{array}{c} 1 \\ 0 \\ 0 \end{array}\right]\right) X' + \left(\mathbf{R}_{\vec{\theta}}\left[\begin{array}{c} 0 \\ 1 \\ 0 \end{array}\right]\right) Y' + \left(\mathbf{R}_{\vec{\theta}}\left[\begin{array}{c} 0 \\ 0 \\ 1 \end{array}\right]\right) Z' \tag{14.28}$$

Tensors in Quantum Mechanics (cont.)

To make this more concrete, let's use the specific example we considered in Section 13.2, in which the F' axes are obtained by rotating the F axes by θ CCW about $\hat{z}$; that is

$$\mathbf{R}_{\vec{\theta}} = \begin{bmatrix} c_\theta & -s_\theta & 0 \\ s_\theta & c_\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{R}_{-\vec{\theta}} = \begin{bmatrix} c_\theta & s_\theta & 0 \\ -s_\theta & c_\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{14.29}$$

($c_\theta = \cos\theta$, $s_\theta = \sin\theta$) The tricky coordinate representations are then

$$\underline{(\vec{R})}' = \begin{bmatrix} X c_\theta + Y s_\theta \\ -X s_\theta + Y c_\theta \\ Z \end{bmatrix} \qquad \underline{(\vec{R}')} = \begin{bmatrix} X' c_\theta - Y' s_\theta \\ X' s_\theta + Y' c_\theta \\ Z' \end{bmatrix} \tag{14.30}$$

Let us compare to the forms we wrote before relating the transformed and untransformed operators:

$$X' = X\cos\theta + Y\sin\theta \qquad Y' = -X\sin\theta + Y\cos\theta \qquad Z' = Z \tag{14.31}$$

It becomes clear that the components of the coordinate representation of the operator $\vec{R}$ in the F' frame are the same as those of $\vec{R}'$ in the F' frame, and similarly the components of the coordinate representation of the operator $\vec{R}'$ in the F frame are the same as those of $\vec{R}$ in the F frame.

Tensors in Quantum Mechanics (cont.)

While our choice of $\widehat{\theta} = \widehat{z}$ was algebraically convenient, it was not fundamental; the same equality would hold in general for any rotation transformation, though it would be far messier to prove it. So, to summarize: let the operators R_i associated with a given coordinate system F define the coordinate representation $(\underline{\vec{R}})$ in F of a vector $\vec{R}$. Let there be another coordinate system F' obtained from F by a CCW rotation by an angle θ around an axis $\widehat{\theta}$. Apply the standard vector transformation law to $\vec{R}$ to obtain its components $(\underline{\vec{R}})'$ in F' from its components $(\underline{\vec{R}})$ in F:

$$(\underline{\vec{R}})' = \mathbf{R}_{-\vec{\theta}}(\underline{\vec{R}}) \tag{14.32}$$

Also, apply a unitary transformation using the quantum mechanical tranformation operator $T(\vec{\theta})$ to each component of $(\underline{\vec{R}})$ to obtain the representation $(\underline{\vec{R}'})'$ of the transformed operator $\vec{R}'$ in F' (this is essentially a definition of $\vec{R}'$):

$$(\underline{\vec{R}'})' \equiv T(\vec{\theta})\,(\underline{\vec{R}})\,T^{\dagger}(\vec{\theta}) \tag{14.33}$$

Then, we have shown that

$$(\underline{\vec{R}'})' = (\underline{\vec{R}})' \tag{14.34}$$

That is, *the untransformed and transformed operators are just different coordinate representations of the same vector operator.*

Our example used the $\vec{R}$ operator, for which we know the quantum mechanical unitary transformation law. The same result would hold for $\vec{P}$ because it has the same unitary transformation law. Since any observable operator must be constructed from $\vec{R}$ and $\vec{P}$, we may thus conclude that any observable vector operator will satisfy the same connection between its quantum mechanical unitary transformation law and its classical transformation law. By extension, we may conclude that the same thing will hold for tensor operators of any rank constructed from $\vec{R}$ and $\vec{P}$. This relationship is wonderful because it shows that our classical definition of the transformation of a tensor from one frame to another is consistent with our component-by-component quantum mechanical unitary transformation law for operators. We are therefore able to define in completely unambiguous fashion tensor operators whose tensor representations are consistent with the classical tensors they are motivated by.

Lecture 37:
Tensors in Classical and Quantum Mechanics

Revision Date: 2008/02/11

Tensors in Quantum Mechanics

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 9.1 to calculate

$$\psi_{q'}(\vec{r}') = \langle\vec{r}'|\psi\rangle = \langle\vec{r}|T^\dagger(\vec{\theta})|\psi\rangle = \langle\vec{r}|\exp\left(\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{L}\right)|\psi\rangle \tag{14.35}$$

$$\psi_q'(\vec{r}) = \langle\vec{r}|\psi'\rangle = \langle\vec{r}|T(\vec{\theta})|\psi\rangle = \langle\vec{r}|\exp\left(-\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{L}\right)|\psi\rangle \tag{14.36}$$

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.

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.

So, let us define a vector state to be an object whose representation in a particular frame F consists of N (scalar) Hilbert space states $\{|\psi_i\rangle\}$, $i = 1, \ldots N$, and whose representation in any other coordinate system F' related to F by a rotation $\vec{\theta}$ is given by

$$
\left[
\begin{array}{c}
\langle \vec{r}' | \psi_1 \rangle \\
\langle \vec{r}' | \psi_2 \rangle \\
\vdots \\
\langle \vec{r}' | \psi_N \rangle
\end{array}
\right]
= \mathbf{R}_{-\vec{\theta}}
\left[
\begin{array}{c}
\langle \vec{r} | T^{\dagger}(\vec{\theta}) | \psi_1 \rangle \\
\langle \vec{r} | T^{\dagger}(\vec{\theta}) | \psi_2 \rangle \\
\vdots \\
\langle \vec{r} | T^{\dagger}(\vec{\theta}) | \psi_N \rangle
\end{array}
\right]
\tag{14.37}
$$

or, equivalently

$$
\langle \vec{r}' | \psi_i \rangle = \sum_{j=1}^{N} \left(\mathbf{R}_{-\vec{\theta}} \right)_{ij} \langle \vec{r} | T^{\dagger}(\vec{\theta}) | \psi_j \rangle
\tag{14.38}
$$

where $T^{\dagger}(\vec{\theta})$ is the operator one would normally apply to perform a passive rotation on the state $|\psi\rangle$ as we did on the last page in calculating $\langle \vec{r}' | \psi \rangle$ and where $\mathbf{R}_{-\vec{\theta}}$ is the rotation matrix one would normally apply to obtain the representation $\underline{v}'$ in F' of a vector $\vec{v}$ from its representation $\underline{v}$ in F. That is, we apply the quantum mechanical unitary transformation operator $T^{\dagger}(\vec{\theta})$ to each component Hilbert space state, then we apply the classical rotation matrix to the ensemble, treating each component Hilbert space wavefunction as the component of a classical vector.

Let us try to generalize this. To avoid being stuck with the position-basis representation obtained by taking the product with $\langle \vec{r} |$ or $\langle \vec{r}\,' |$, though, we must write our definition for an *active* rotation transformation — *i.e.*, physically rotating the state relative to the underlying space. We define a tensor state of rank *n* as an object that consists of a set of N^n (scalar) Hilbert space states $\{ |\psi_{i_1 \cdots i_n} \rangle \}$, $i_k = 1, \ldots, N$, $k = 1, \ldots, n$ that are transformed by an active rotation transformation $T(\vec{\theta})$ by

$$|\psi'_{i_1 \cdots i_n} \rangle = \sum_{j_1, \ldots, j_n = 1}^{N} \left(\mathbf{R}_{\vec{\theta}} \right)_{i_1 j_1} \cdots \left(\mathbf{R}_{\vec{\theta}} \right)_{i_n j_n} T(\vec{\theta}) |\psi_{j_1 \cdots j_n} \rangle \qquad (14.39)$$

Let us explain in detail why we needed to consider an active transformation. The problem is that a passive transformation really can only be defined for the position-basis wavefunction – since there is supposed to be no change in the state, just a change in the coordinate system, one must always speak of how the projection onto the position basis changes under a passive transformation. That restriction will prove inconvenient in the future. So, using the known relationship between passive and active transformations, we postulate the above definition for the behavior under active transformations. An active transformation by $\vec{\theta}$ of course looks like a passive transformation by $-\vec{\theta}$ when the position-basis projections are taken; but, an active transformation creates a new state $|\psi'_{i_1 \cdots i_n} \rangle$, which is what we need in order to free the definition of a tensor state from the position basis.

Tensor Particle States — Formal Definition

The above factorization of the transformation law into the Hilbert space transformation operator $T(\vec{\theta})$ and the classical transformation operator $\mathbf{R}_{-\vec{\theta}}$ suggests that the natural way to write these tensor states is as direct products. Let's first begin with vector states. Suppose $\mathcal{V}$ is the space of scalar Hilbert space states, with basis $\{|\phi_i\rangle\}$. Let $\tau^{(1)}$ denote the vector space of vectors for spatial dimension $N = 3$. A valid basis for $\tau^{(1)}$ is the set of unit vectors $\{\vec{e}_i\}$ of the frame F. A basis for the product space $\mathcal{V} \otimes \tau^{(1)}$ consists of direct products of these two bases, $\{|\phi_i\rangle \otimes \vec{e}_j\}$. Any vector state may be expanded in this basis via

$$|\vec{\psi}\rangle = \sum_i \sum_{j=1}^{N} C_{ij}\left(|\phi_i\rangle \otimes \vec{e}_j\right) \tag{14.40}$$

where the vector symbol indicates that $|\vec{\psi}\rangle$ is a vector state. To match up with the notation used before, we treat $\otimes$ as multiplication of scalar numbers against scalar Hilbert space states, so the above may be written in F as

$$|\vec{\psi}\rangle = \sum_i \left(C_{i,1}\left[\begin{array}{c} |\phi_i\rangle \\ 0 \\ 0 \end{array}\right] + C_{i,2}\left[\begin{array}{c} 0 \\ |\phi_i\rangle \\ 0 \end{array}\right] + C_{i,3}\left[\begin{array}{c} 0 \\ 0 \\ |\phi_i\rangle \end{array}\right] \right) \tag{14.41}$$

Written in terms of vector space direct products, it becomes clear that the transformation operator on such direct product states is just the direct product of the transformation operators for each individual space. That is, we know the transformation operator in $\mathcal{V}$ is $T(\vec{\theta})$ and the transformation operator in $\tau^{(1)}$ is $\mathbf{R}_{\vec{\theta}}$, so we expect that the transformation operator for the direct product space ought to be

$$\mathbf{T}(\vec{\theta}) = T(\vec{\theta}) \otimes \mathbf{R}_{\vec{\theta}} \tag{14.42}$$

so, when we act on a particular coordinate representation, we obtain

$$|\vec{\psi}'\,\rangle = \mathbf{T}(\vec{\theta})|\vec{\psi}\,\rangle \tag{14.43}$$

$$= \mathbf{T}(\vec{\theta}) \sum_i \sum_{j=1}^{N} C_{ij} \left(|\phi_i\,\rangle \otimes \vec{e}_j \right) \tag{14.44}$$

$$= \sum_i \sum_{j=1}^{N} C_{ij} \left(\left[T(\vec{\theta})|\phi_i\,\rangle \right] \otimes \left[\mathbf{R}_{\vec{\theta}}\,(\underline{\vec{e}_j}) \right] \right) \tag{14.45}$$

The natural extension to tensor states is obvious. Let $\tau^{(n)}$ be the vector space of rank n tensors and let $\mathcal{V}$ be the space of scalar Hilbert space states as before. Then, our vector space of tensor states is

$$V = \mathcal{V} \otimes \tau^{(n)} \tag{14.46}$$

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 $\mathcal{E}_{i_1 \cdots i_n}$, $i_k = 1, \ldots, N$, $k = 1, \ldots, n$, where the indices indicate which entry is nonzero in the coordinate representation in F.

Tensors in Quantum Mechanics (cont.)

The above basis for $\tau^{(n)}$ suggests a more fundamental way of writing $\tau^{(n)}$ and its basis. Each basis element $\mathcal{E}_{i_1 \cdots i_n}$ can be written as an outer product of unit vectors:

$$\mathcal{E}_{i_1 \cdots i_n} = \vec{e}_{i_1} \cdots \vec{e}_{i_n} \tag{14.47}$$

It should be clear that an outer product of vectors is just a kind of direct product. That is, one could write the above as

$$\mathcal{E}_{i_1 \cdots i_n} = \vec{e}_{i_1} \otimes \cdots \otimes \vec{e}_{i_n} = \prod_{k=1}^{n} \vec{e}_{i_k} \tag{14.48}$$

and thus the vector space of tensors of rank n is just a direct product of vector spaces of vectors:

$$\tau^{(n)} = \tau^{(1)} \otimes \cdots \otimes \tau^{(1)} = \prod_{k=1}^{n} \tau^{(1)} \tag{14.49}$$

where there are n elements in the direct product. 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.

Therefore, our vector space of tensor states is

$$V = \mathcal{V} \otimes \prod_{k=1}^{n} \tau^{(1)} \tag{14.50}$$

This space is a Hilbert space because it inherits the necessary properties from the factor spaces. Any state can be expanded as

$$|\psi^{(n)}\rangle = \sum_{i} \sum_{i_1,\ldots,i_n=1}^{N} C_{i\,i_1 \cdot i_n} \left(|\phi_i\rangle \otimes \prod_{k=1}^{n} \vec{e}_{i_k} \right) \tag{14.51}$$

where the $\psi^{(n)}$ symbol is meant to indicate that the state is a tensor state of rank n as opposed to a scalar state. A vector state may be written as $|\psi^{(1)}\rangle$ or as $|\vec{\psi}\rangle$.

Tensors in Quantum Mechanics (cont.)

The natural rotation operator is obvious the tensor product of rotation operators in the factor spaces:

$$\mathbf{T}^{(n)}(\vec{\theta}) = T(\vec{\theta}) \otimes \mathbf{R}_{\vec{\theta}} \otimes \cdots \otimes \mathbf{R}_{\vec{\theta}} = T(\vec{\theta}) \otimes \prod_{k=1}^{n} \mathbf{R}_{\vec{\theta}} \tag{14.52}$$

where the $^{(n)}$ superscript indicates that the "tensor" part of $\mathbf{T}^{(n)}(\vec{\theta})$ works on rank n tensors. The action of $\mathbf{T}^{(n)}(\vec{\theta})$ is obviously

$$|\psi^{(n)\,\prime}\rangle = \mathbf{T}^{(n)}(\vec{\theta})|\psi^{(n)}\rangle \tag{14.53}$$

$$= \mathbf{T}^{(n)}(\vec{\theta}) \sum_{i} \sum_{i_1,\ldots,i_n=1}^{N} C_{i\,i_1\cdots i_n} \left(|\phi_i\rangle \otimes \prod_{k=1}^{n} \vec{e}_{i_k} \right) \tag{14.54}$$

$$= \sum_{i} \sum_{i_1,\ldots,i_n=1}^{N} C_{i\,i_1\cdots i_n} \left(\left[T(\vec{\theta})|\phi_i\rangle \right] \otimes \prod_{k=1}^{n} \left[\mathbf{R}_{\vec{\theta}}\vec{e}_{i_k} \right] \right) \tag{14.55}$$

which performs the desired active rotation on each factor of the direct product state.

Tensors in Quantum Mechanics (cont.)

Angular Momentum Operators for Tensor Particle States

We now focus on defining the correct operators for these tensor particle states. We will begin with angular momentum, which we will find will provide some physical interpretation and allows us to define other kinds of operators sensibly. The reason we begin with angular momentum is that angular momentum operators are obviously tied to the rotation operators that we use to *define* tensor states.

Recall that we showed in Section 13.2 that the generator of rotations about the z axis is L_z by calculating the action of an infinitesimal rotation and relating it to the quantum mechanical analogue of the classical l_z angular momentum. Let's use the same procedure here, though done less pedantically since we are now experts at doing this kind of thing.

An infinitesimal rotation operator $\mathbf{T}^{(n)}(\delta\vec{\theta})$ for tensor states of rank n is

$$\mathbf{T}^{(n)}(\delta\vec{\theta}) = T(\delta\vec{\theta}) \otimes \prod_{k=1}^{n} \mathbf{R}_{\delta\vec{\theta}} = \left(I - \frac{i}{\hbar}\, \delta\vec{\theta} \cdot \vec{L} \right) \otimes \prod_{k=1}^{n} \left(I + \delta\vec{\theta} \cdot \vec{\mathbf{M}} \right) \qquad (14.56)$$

Let's allow this to act on a tensor state:

$$|\psi^{(n)\,\prime}\,\rangle = \mathbf{T}^{(n)}(\delta\vec{\theta})|\psi^{(n)}\,\rangle \tag{14.57}$$

$$= \sum_i \sum_{i_1,\ldots,i_n=1}^{N} c_{i\,i_1\cdots i_n} \left(\left[\left(I - \frac{i}{\hbar}\,\delta\vec{\theta}\cdot\vec{L} \right) |\phi_i\,\rangle \right] \otimes \prod_{k=1}^{n} \left[\left(I + \delta\vec{\theta}\cdot\vec{\mathbf{M}} \right) \vec{e}_{i_k} \right] \right) \tag{14.58}$$

Since this is an infinitesimal transformation, we want to keep the zeroth order term (proportional to $|\psi^{(n)}\,\rangle$) and terms of first order in $\delta\vec{\theta}$. It should be clear that there will be one term for each factor, in which that factor's $\delta\vec{\theta}$ term multiplies the identity operator for all the other factors. Let's write it out for $n = 2$ for the sake of being explicit, and this will then motivate the generic formula:

$$\left[\left(I - \frac{i}{\hbar}\,\delta\vec{\theta}\cdot\vec{L} \right) |\phi_i\,\rangle \right] \otimes \left[\left(I + \delta\vec{\theta}\cdot\vec{\mathbf{M}} \right) \vec{e}_{i_1} \right] \otimes \left[\left(I + \delta\vec{\theta}\cdot\vec{\mathbf{M}} \right) \vec{e}_{i_2} \right] \tag{14.59}$$

$$= I\otimes I\otimes I \left(-\frac{i}{\hbar}\,\delta\vec{\theta}\cdot\vec{L} \right) \otimes I\otimes I + I\otimes \delta\vec{\theta}\cdot\vec{\mathbf{M}}\otimes I + I\otimes I\otimes \delta\vec{\theta}\cdot\vec{\mathbf{M}} \tag{14.60}$$

$$= I\otimes I\otimes I - \frac{i}{\hbar}\,\delta\vec{\theta}\cdot\left(\vec{L}\otimes I\otimes I + I\otimes i\,\hbar\vec{\mathbf{M}}\otimes I + I\otimes I\otimes i\,\hbar\vec{\mathbf{M}} \right) \tag{14.61}$$

Tensors in Quantum Mechanics (cont.)

Clearly, then, we may write the action of the infinitesmial rotation as

$$\mathbf{T}^{(n)}(\delta\vec{\theta})|\psi^{(n)}\rangle = |\psi^{(n)}\rangle - \frac{i}{\hbar}\,\delta\vec{\theta}\cdot\left[\vec{L}\otimes\prod_{i=1}^{n}{}^{(i)}I + \sum_{i=I}^{n}I\otimes i\,\hbar\,{}^{(i)}\vec{\mathbf{M}}\otimes\prod_{k\neq i}^{N}{}^{(k)}I\right]|\psi^{(n)}\rangle \tag{14.62}$$

where the ${}^{(k)}$ subscripts in front of an operator indicate which of the $n\,\tau^{(1)}$ product spaces the operator belongs to (because they all identical). It becomes clear that we may define a generator

$$\vec{\mathcal{J}} = \vec{L}\otimes\prod_{i=1}^{n}{}^{(i)}I + \sum_{i=1}^{n}I\otimes i\,\hbar\,{}^{(i)}\vec{\mathbf{M}}\otimes\prod_{k\neq i}^{N}{}^{(k)}I \tag{14.63}$$

and that the infinitesimal rotation operator is

$$\mathbf{T}^{(n)}(\delta\vec{\theta}) = I - \frac{i}{\hbar}\,\delta\vec{\theta}\cdot\vec{\mathcal{J}} \tag{14.64}$$

A finite rotation operator will then be, by the usual integration argument,

$$\mathbf{T}^{(n)}(\vec{\theta}) = \exp\left(-\frac{i}{\hbar}\,\delta\vec{\theta}\cdot\vec{\mathcal{J}}\right) \tag{14.65}$$

Since $\vec{\mathcal{J}}$ generates rotations in the Hilbert space of tensor states, it is thus natural to identify $\vec{\mathcal{J}}$ as the angular momentum operator in the Hilbert space of tensor states (which explains why we called it $\vec{\mathcal{J}}$). It is straightforward to check that it satisfies the standard angular momentum operator commutation relations

$$[\mathcal{J}_i, \mathcal{J}_j] = \epsilon_{ijk}\,i\,\hbar\,J_k \qquad \Longleftrightarrow \qquad \vec{\mathcal{J}}\times\vec{\mathcal{J}} = i\,\hbar\,\vec{\mathcal{J}} \tag{14.66}$$

because each term in the sum defining $\vec{\mathcal{J}}$ separately satisfies these commutation relations. Thus, all our results derived based on these relations — especially regarding the eigenvalues-eigenvector structure of $\vec{\mathcal{J}}$ and $\mathcal{J}^2 = \vec{\mathcal{J}}\cdot\vec{\mathcal{J}}$ — will hold.

We thus see why we foreshadowed earlier that the $\mathbf{\vec{M}}$ matrix would become interesting for discussing particle spin — it contributes a piece to the total angular momentum operator $\vec{\mathcal{J}}$ for these tensor states.

The Eigenvector-Eigenvalue Problem of $i\hbar\vec{\mathbf{M}}$

Before we go on to what the eigenvalues and eigenvectors of $\vec{\mathcal{J}}$ (actually, $\mathcal{J}_z$ and $\mathcal{J}^2$) are, we clearly need to sort out what the eigenvalues and eigenvalues of $i\hbar\vec{\mathbf{M}}$ in $\tau^{(1)}$ are. As noted above, $i\hbar\vec{\mathbf{M}}$ satisfies the standard angular momentum commutation relations. Our work in the $|j,m\rangle$ basis showed that any set of operators that satisfies these relations yields a set of eigenstates $|j,m\rangle$ of $\mathbf{M}^2$ and $\mathbf{M}_z$ with allowed eigenvalues defined by

$$\left(i\hbar\vec{\mathbf{M}}\right)^2|j,m\rangle = \hbar^2 j(j+1)|j,m\rangle \qquad i\hbar\mathbf{M}_z|j,m\rangle = \hbar m|j,m\rangle \tag{14.67}$$

$$j = \frac{k}{2} \quad k\text{ any integer} \quad m = -j, -(j-1), \cdots, j-1, j \tag{14.68}$$

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 scalar states 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 $|j,m\rangle$ states exist in $\tau^{(1)}$?

Tensors in Quantum Mechanics (cont.)

Well, just as we did for orbital angular momentum, let's just write down the specific eigenvector-eigenvalue problem. First, we remind you:

$$i \hbar \mathbf{M}_z = i \hbar \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \left(i \hbar \vec{\mathbf{M}} \right)^2 = 2 \hbar^2 \, \mathbf{I} \qquad (14.69)$$

From the form of $(i \hbar \vec{\mathbf{M}})^2$, we immediately see $j = 1$ is the only allowed value because the eigenvalues of any J^2-type operator are $j(j+1) \hbar^2$. We can see quickly that the expected values m, $m = -1, 0, 1$, are indeed realized by finding the roots of the characteristic polynomial for the $i \hbar \mathbf{M}_z$ eigenvalue problem:

$$m \left(m^2 - \hbar^2 \right) = 0 \qquad \implies \qquad m = 0, \pm 1 \qquad (14.70)$$

That is, vectors in 3 spatial dimensions are $j = 1$ objects. For completeness, one can show the eigenvectors are

$$\vec{v}_{+1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \\ 0 \end{bmatrix} \qquad \vec{v}_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad \vec{v}_{-1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \\ 0 \end{bmatrix} \qquad (14.71)$$

The operators we used above are complex, as are two of the eigenvectors, so we really ought to expand $\tau^{(1)}$ to allow complex coefficients. There are two reasons to just go ahead and do this rather than worrying unduly about the fact that classical vectors are not complex. First, 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. Second, when we form linear combinations of states in the direct product space with the scalar Hilbert space states, we will allow complex coefficients. Since we can create complex linear combinations by doing it in the product space, so we might as well allow them in the factor space to make our lives easier. For example, even if $\tau^{(1)}$ is restricted to real vectors, we can construct something in the product space that is basically $|\psi\rangle \otimes \vec{v}_{\pm 1}$ via

$$|\psi, m = \pm 1\rangle = \frac{1}{\sqrt{2}} \left[|\psi\rangle \otimes \vec{e}_1 \pm i |\psi\rangle \otimes \vec{e}_2 \right] \tag{14.72}$$

It would be much more convenient to be able to write this as

$$|\psi, m = \pm 1\rangle = |\psi\rangle \otimes \vec{v}_{\pm 1} \tag{14.73}$$

which we can do if we expand $\tau^{(1)}$ as desired.

Another issue is: How does this relate to the $\bar{J}^{(1)}$ matrices obtained in our discussion of the $|j, m\rangle$ basis? Well, we simply need to apply a unitary transformation from the conventional $\{\vec{e}_i\}$ basis for $\tau^{(1)}$ to the basis of the above eigenvectors $\{\vec{v}_i\}$. In the latter basis, the $\bar{J}^{(1)}$ are the matrix representations of $(i\,\hbar\,\vec{\mathbf{M}})$ (as one can verify by explicitly doing the unitary transformation).

Note: Be sure not to confuse the unitary transformation necessary for diagonalizing $i\,\hbar\,\mathbf{M}_z$ with a spatial rotation. Because $\vec{v}_{\pm 1}$ are complex, one can never obtain $\vec{v}_{\pm 1}$ from the $\{\vec{e}_i\}$ by an orthogonal spatial rotation transformation: the matrix representations of such a rotation only contains real numbers! The spatial rotations are a subset of all unitary transformations on $\tau^{(1)}$, and the necessary unitary transformation is outside of that subset.

Higher Integral j Values?

We have been told that integral values of j for $j \geq 2$ are allowed by the commutation relations for angular momentum operators. Yet we have found no manifestation of them yet. Where are they?

We shall see later when we discuss addition of angular momentum that the tensor states provide the higher integer-j states. In particular, we shall see that tensor states of rank n contain states with all integer values of j from 0 to n. But, because the natural way to build up tensor states is via direct products of vector states, which result in addition of the angular momentum operators in the factor spaces, let us defer that discussion until we come to addition of angular momentum. We could also define them by fiat, as we will do with half-integer j states below, but that is unsatisfying by comparison to the physically motivated discussion we can have in terms of tensor states of rank n.

Half-Integral Values of *j*

What about half-integral values of *j*? Let's begin with the lowest possible value, $j = 1/2$, first, and then work from there. Since we have no prior classical motivation to begin with for half-integral *j* values, the obvious thing to do is simply use the $\bar{J}^{(k/2)}$ matrices defined earlier as the matrix representations of the desired operators. They are

$$\mathbf{J}_x^{(1/2)} = \frac{\hbar}{2} \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right] \qquad \mathbf{J}_y^{(1/2)} = \frac{\hbar}{2} \left[\begin{array}{cc} 0 & -i \\ i & 0 \end{array} \right] \qquad \mathbf{J}_z^{(1/2)} = \frac{\hbar}{2} \left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right]$$
(14.74)

$$\left(\mathbf{J}^2 \right)^{(1/2)} = \frac{3}{4} \hbar^2 \, \mathbf{I}$$
(14.75)

The eigenvectors are trivial because $(\mathbf{J}^2)^{(1/2)}$ and $\mathbf{J}_z$ are already diagonal:

$$\vec{v}_{+1/2} = \left[\begin{array}{c} 1 \\ 0 \end{array} \right] \qquad \vec{v}_{-1/2} = \left[\begin{array}{c} 0 \\ 1 \end{array} \right]$$
(14.76)

The Hilbert space in which these operators work is the space of all complex two-element vectors, consisting of linear combinations with complex coefficients of the above two eigenvectors. Let us call this space $\tau^{(1/2)}$ just as the space of all three-element vectors with complex coefficients was termed $\tau^{(1)}$, and we call the 2-element column vectors spinors. We point out that $\mathbf{M}_z$ was not automatically diagonalized because its form was specified ahead of time by requiring that it generate classical rotations about the z axis. Since we are pulling $\vec{\mathbf{J}}^{(1/2)}$ out of thin air, we are free to choose whatever basis vectors we want for the space, so we choose the eigenvectors of $\mathbf{J}_z^{(1/2)}$.

Hilbert spaces for higher half-integral values of j may be done by fiat in a similar fashion; obviously,

$$\left(\mathbf{J}^2\right)^{(k/2)} = j(j+1)\,\hbar^2\,\mathbf{I} \qquad \mathbf{J}_z^{(k/2)} = \mathrm{diag}(j, j-1, \ldots, -(j-1), -j) \qquad (14.77)$$

and the matrix representations of $J_{x,y}^{(k/2)}$ were given in detail before and in Shankar. However, we note that we can also obtain all of these by direct products of $\tau^{(1/2)}$ with $\tau^{(n)}$ spaces, so we will return to them when we consider addition of angular momentum.

Lecture 38:
Kinematics and Dynamics of Spin-1/2 Systems

Revision Date: 2008/02/01

Kinematics of Spin-1/2 Systems

Motivation

Let's now specialize to considering spin-1/2 systems. The reason we spend so much time on this particular example is:

- ▶ All the fundamental stable matter particles are spin-1/2.
- ▶ Most importantly, the electron is spin-1/2, so many aspects of atomic theory, which was the first good laboratory for quantum mechanics, depend on includiing the spin-1/2 nature of the electron.
- ▶ It's simple — there are only two eigenstates $m = \pm 1/2$ and all the operators acting on the $\tau^{(1/2)}$ space can written as linear combination of the three components of $\vec{\mathbf{J}}^{(1/2)}$ and the identity matrix. Even rotation transformations have a simple form.
- ▶ The system is remarkably physically rich in spite of the fact it has only two degrees of freedom.

Algebraic Preliminaries

Let's denote $\vec{J}^{(1/2)}$ by $\vec{S}$ for brevity, and we also write $\vec{S} = (\hbar/2)\,\vec{\sigma}$ where $\vec{\sigma}$ are the Pauli matrices. We list some useful algebraic properties of $\vec{\sigma}$ for future use:

$$\sigma_x = \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right] \qquad \sigma_y = \left[\begin{array}{cc} 0 & -i \\ i & 0 \end{array} \right] \qquad \sigma_z = \left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right] \tag{14.78}$$

$$[\sigma_i, \sigma_j] = 2\,i\,\epsilon_{ijk}\,\sigma_k \qquad [\sigma_i, \sigma_j]_+ = 2\,\delta_{ij}\,\mathbf{I} \tag{14.79}$$

$$\sigma_i\sigma_j = -\sigma_j\sigma_i \qquad \sigma_i\sigma_j = i\,\epsilon_{ijk}\,\sigma_k \tag{14.80}$$

$$\sigma_i^2 = \mathbf{I} \qquad (\hat{n}\cdot\vec{\sigma})^2 = \mathbf{I} \qquad \vec{\sigma}\cdot\vec{\sigma} = 3\,\mathbf{I} \tag{14.81}$$

$$\text{Tr}\,\sigma_i = 0 \tag{14.82}$$

$$(\vec{a}\cdot\vec{\sigma})(\vec{b}\cdot\vec{\sigma}) = \vec{a}\cdot\vec{b}\,\mathbf{I} + i\left(\vec{a}\times\vec{b}\right)\cdot\vec{\sigma} \tag{14.83}$$

$$\text{Tr}\,\sigma_i\sigma_j = 2\,\delta_{ij} \tag{14.84}$$

The $\vec{\sigma}$ along with **I** form a complete basis for the vector space of 2×2 complex matrices (if complex coefficients are allowed in making linear combinations) and for the vector space of 2×2 Hermitian matrices (if only real coefficients are allowed). We define $\sigma_0 = $ **I** for brevity. The space is an inner product space via definition of the inner product on the basis elements by

$$\langle \sigma_i \, | \, \sigma_j \rangle = \frac{1}{2} \, \text{Tr} \, \sigma_i \sigma_j \tag{14.85}$$

For the sake of brevity, we will indicate eigenstates of $\widehat{n} \cdot \vec{\mathbf{S}}$ by

$$|\widehat{n}, + \rangle \qquad |\widehat{n}, - \rangle \tag{14.86}$$

and, since it is our default,

$$|+\rangle \qquad |-\rangle \tag{14.87}$$

will imply $\widehat{n} = \widehat{z}$.

Rotation Transformation Operators

The above simple form for the $\vec{\mathbf{S}}$ matrices implies that writing out a matrix representation of the rotation transformation operator in $\tau^{(1/2)}$ is easy. We denote this rotation operator by $\mathbf{R}_{\vec{\theta}}^{(1/2)}$ for sake of consistency with our usage of $\mathbf{R}_{\vec{\theta}}$ for vectors (and tensors). So

$$\mathbf{R}_{\vec{\theta}}^{(1/2)} = \exp\left(-\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{\mathbf{S}}\right) = \exp\left(-\frac{i}{2}\,\theta\,\widehat{\theta}\cdot\vec{\boldsymbol{\sigma}}\right) \tag{14.88}$$

$$= \sum_{n=0}^{\infty}\left(-\frac{i}{2}\,\theta\right)^{n}\frac{\left(\widehat{\theta}\cdot\vec{\boldsymbol{\sigma}}\right)^{n}}{n!} \tag{14.89}$$

$$= \sum_{n=0}^{\infty}\left[\mathbf{I}\frac{1}{(2n)!}\left(-\frac{i}{2}\,\theta\right)^{2n} + \widehat{\theta}\cdot\vec{\boldsymbol{\sigma}}\frac{1}{(2n+1)!}\left(-\frac{i}{2}\,\theta\right)^{2n+1}\right] \tag{14.90}$$

$$= \mathbf{I}\cos\left(\frac{\theta}{2}\right) - \widehat{\theta}\cdot\vec{\boldsymbol{\sigma}}\,i\,\sin\left(\frac{\theta}{2}\right) \tag{14.91}$$

Let us try using this form to rotate the state $|\hat{z}, +\rangle$ to the state $|\hat{x}, +\rangle$. It is a reasonable guess to think that the rotation angle has to be $\theta = \pi/2$ and that the rotation axis has to be $\hat{y}$. (Remember, $\theta > 0$ corresponds to CCW rotation.) Let's try that:

$$\mathbf{R}_{\frac{\pi}{2}\,\hat{x}}^{(1/2)} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \left[\mathbf{I} \cos \frac{\pi}{4} - \boldsymbol{\sigma}_y \, i \, \sin \frac{\pi}{4} \right] \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{14.92}$$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{14.93}$$

Let's check if this is indeed the $+\hbar/2$ eigenvector of $\mathbf{S}_x$:

$$\mathbf{S}_x \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{\hbar}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{\hbar}{2} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{14.94}$$

So, indeed, we get what we expected. More generally, if we rotate $|\hat{z}, +\rangle$ by spherical coordinate angles (θ, ϕ), so that is will be an eigenstate of $\hat{n} \cdot \vec{\mathbf{S}}$ where $\hat{n}$ points along (θ, ϕ), we find the result is

$$|\hat{n}, +\rangle = \mathbf{R}_{\phi\,\hat{z}}^{(1/2)} \mathbf{R}_{\theta\,\hat{y}}^{(1/2)} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} e^{-i\phi/2} \cos \frac{\theta}{2} \\ e^{i\phi/2} \sin \frac{\theta}{2} \end{bmatrix} \tag{14.95}$$

Dynamics of Spin-1/2 Systems

Overview

We will proceed as follows in trying to understand the quantum dynamics of spin-1/2 systems:

▶ Review the classical mechanics of an orbital magnetic moment in a magnetic field.

▶ Study the quantum mechanics of an orbital magnetic moment in a magnetic field based on what we already know about quantum mechanics.

▶ Use the above to motivate the Hamiltonian for a spin-1/2 magnetic moment in a magnetic field and study its dynamics.

▶ Study the specific application of quantum mechanical paramagnetic resonance, which requires only the spin Hamiltonian, including a discussion of its application to a search for the neutron electric dipole moment.

▶ Consider problems involving orbital and spin pieces, such as the Zeeman effect in the SHO.

Classical Mechanics of a Magnetic Moment in a Magnetic Field

Consider a classical magnetic moment $\vec{\mu}$, which you can think of as a current loop with current I and area A such that $\mu = IA/c$ and $\widehat{n} = \vec{\mu}/\mu$ is the normal to the loop such that the current flows CCW around $\widehat{n}$. You know from classical mechanics that, in a uniform magnetic field $\vec{B}$, such a magnetic moment experiences a torque

$$\vec{T} = \vec{\mu} \times \vec{B} \tag{14.96}$$

The torque tries to orient $\vec{\mu}$ along $\vec{B}$. Integrating the torque to obtain a potential energy yields

$$U(\theta) = -\vec{\mu} \cdot \vec{B} \tag{14.97}$$

Now, if instead of a current loop, we have a particle moving in a circular orbit of radius r, we may calculate its *orbital* magnetic moment to be

$$\vec{\mu} = \frac{IA}{c} \,\widehat{n} = \frac{1}{c} \left(\frac{q\,v}{2\,\pi\,r} \right) \left(\pi\,r^2 \right) \widehat{n} = \frac{q}{2\,m\,c}\, m\,v\,r\,\widehat{n} = \frac{q}{2\,m\,c}\, \vec{L} \qquad (14.98)$$

where $\vec{L}$ is the angular momentum of the particle. We define the gyromagnetic ratio,

$$\gamma = \frac{|\vec{\mu}|}{|\vec{L}|} = \frac{q}{2\,m\,c} \qquad (14.99)$$

When we consider the moment to arise from an angular momentum, the dynamics change: the equation of motion changes from

$$\frac{d}{dt}\,\vec{L} = \vec{\mu} \times \vec{B} \qquad (14.100)$$

with no $\vec{L}$ on the right side to

$$\frac{d}{dt}\,\vec{L} = \gamma\,\vec{L} \times \vec{B} \qquad (14.101)$$

where $\vec{L}$ appears on the right side.

Dynamics of Spin-1/2 Systems (cont.)

The solution changes from oscillatory motion in θ, the angle between $\vec{\mu}$ and $\vec{B}$, to precession about $\vec{B}$. This can be seen from: 1) $d\vec{L}/dt$ is perpendicular to $\vec{L}$, so the magnitude of $\vec{L}$ does not change; and 2) $\vec{L} \cdot \vec{B}$ is constant:

$$\frac{d}{dt}\left(\vec{L} \cdot \vec{B}\right) = \left(\gamma\,\vec{L} \times \vec{B}\right) \cdot \vec{B} = \gamma\,\vec{L} \cdot \left(\vec{B} \times \vec{B}\right) = 0 \tag{14.102}$$

The precession frequency can be found be breaking $\vec{L}$ up into a piece normal to $\vec{B}$, $\vec{L}_\perp$, and a piece parallel to $\vec{B}$, $\vec{L}_{||}$. The latter experiences no torque because $\vec{L}_{||} \times \vec{B} = 0$. The remaining part precesses in the plane formed by itself and $\vec{T}$ because $\vec{L}_\perp$, $\vec{B}$, and $\vec{T}$ are mutually perpendicular. The torque satisfies

$$\left|\frac{d}{dt}\,\vec{L}_\perp\right| = \gamma\left|\vec{L}_\perp \times \vec{B}\right| = \gamma\left|\vec{L}_\perp\right|\left|\vec{B}\right| \tag{14.103}$$

so the rate of change of the azimuthal angle is directly related to the torque; we can make it a vector by sorting out the signs on the cross-product, giving:

$$\omega_0 = \frac{1}{|\vec{L}_\perp|}\left|\frac{d}{dt}\,\vec{L}_\perp\right| = \gamma\left|\vec{B}\right| \qquad \vec{\omega}_0 = -\gamma\,\vec{B} \tag{14.104}$$

Dynamics of Spin-1/2 Systems (cont.)

Quantum Mechanics of an Orbital Magnetic Moment

Now, let's figure out how to do this problem in quantum mechanics, limiting ourselves for now to the case of a magnetic moment generated by orbital angular momentum, which we can understand in terms of our simple scalar Hilbert space quantum mechanics.

We will state without proof (for now) that the Hamiltonian for the interaction of a particle with magnetic field is

$$ H = \frac{1}{2\,m} \left(\vec{P} - \frac{q}{c}\,\vec{A} \right)^2 = \frac{1}{2\,m}\,\vec{P}^2 - \frac{q}{2\,m\,c}\left(\vec{P} \cdot \vec{A} + \vec{A} \cdot \vec{P} \right) + \frac{q^2}{2\,m\,c^2}\,\vec{A}^2 \quad (14.105) $$

where $\vec{A}$ is the vector potential, $\vec{B} = \vec{\nabla} \times \vec{A}$. For a uniform magnetic field along the $\hat{z}$ direction, you know that

$$ \vec{A} = \frac{B}{2}\left(-y\,\hat{x} + x\,\hat{y} \right) \quad (14.106) $$

which yields $\vec{B} = \vec{\nabla} \times \vec{A} = B\,\hat{z}$.

We assume B/c is small enough compared to the the momenta involved that we may ignore the term quadratic in $\vec{A}$. Let's work on $\vec{P} \cdot \vec{A}$:

$$\vec{P} \cdot \vec{A} = \sum_i P_i A_i(\vec{R}) = \sum_i A_i(\vec{R}) P_i + \sum_i \left[P_i, A_i(\vec{R}) \right] \tag{14.107}$$

We know from previous work that, if $[C, [C, D]] = 0$ and $[D, [C, D]] = 0$, then $[C, f(D)] = \frac{df(D)}{dD} [C, D]$. Here, $C = P_i$, $D = R_j$, and $f(D) = A_i(R_1, R_2, R_3)$. $[P_i, R_j] = -i \hbar \delta_{ij}$, so the conditions $[C, [C, D]] = 0$ and $[D, [C, D]] = 0$ are met. Therefore,

$$\vec{P} \cdot \vec{A} = \sum_i A_i(\vec{R}) P_i - i \hbar \left[\sum_i \frac{\partial A_i}{\partial X_i} \right] \tag{14.108}$$

The operator in brackets is the quantum mechanical equivalent of $\vec{\nabla} \cdot \vec{A}$, which is zero here (and can be made so in general). So we have

$$\vec{P} \cdot \vec{A} = \vec{A} \cdot \vec{P} \tag{14.109}$$

Therefore, the term in the Hamiltonian describing the interaction with the magnetic field is

$$H_{int} = -\frac{q}{2\,m\,c}\left(2\,\vec{A}\cdot\vec{P}\right) \tag{14.110}$$

$$= -\frac{q}{2\,m\,c}\,B(-Y\,P_x + X\,P_y) \tag{14.111}$$

$$= -\frac{q}{2\,m\,c}\,B\,L_z \tag{14.112}$$

The choice of direction of $\vec{B}$ is arbitrary, so it must hold in general

$$H_{int} = -\frac{q}{2\,m\,c}\,\vec{B}\cdot\vec{L} = -\vec{\mu}\cdot\vec{B} \tag{14.113}$$

(We may commute $\vec{B}$ and $\vec{L}$ because $\vec{B}$ is position-independent.) That is, the contribution to the Hamiltonian is what we expect from the classical potential energy, though we must remember that $\vec{\mu}$ is now proportional to the quantum mechanical operator $\vec{L}$.

One can verify using Ehrenfest's Theorem that $\langle\vec{L}\rangle$ precesses around $\vec{B}$. This is slightly nontrivial because it involves calculating $[\vec{L}, H]$.

Quantum Mechanics of a Spin Magnetic Moment

Can we use the above result for the Hamiltonian for the interaction of an orbital magnetic moment with a uniform magnetic field to motivate a Hamiltonian for the interaction of a spin magnetic moment?

As we have said before, $\vec{\mathbf{S}}$ completely spans the space of Hermitian operators that can act on states in $\tau^{(1/2)}$. Therefore, the only choice for a $\vec{\mu}$ operator acting on $\tau^{(1/2)}$ is to make it proportional to $\vec{\mathbf{S}}$. Dimensional analysis and analogy to the orbital case indicates that the form should be

$$\vec{\mu} = g \frac{q}{2\,m\,c} \vec{\mathbf{S}} \tag{14.114}$$

where g is a constant to be determined experimentally and $q = -e$ for an electron. The analogous Hamiltonian is therefore

$$\mathbf{H}_{int} = -\vec{\mu} \cdot \vec{B} = -\frac{g\,q}{2\,m\,c} \vec{\mathbf{S}} \cdot \vec{B} \equiv -\gamma\,\vec{\mathbf{S}} \cdot \vec{B} \tag{14.115}$$

where we write the Hamiltonian as a matrix $\mathbf{H}$ to emphasize that it acts on the 2-element column vectors that make up $\tau^{(1/2)}$ (spinors).

Shankar goes through an interesting digression on the values of g for electrons, protons, and neutrons, and how they all receive significant corrections from perturbative effects associated with the quantum field theory effects of the electromagnetic and strong interactions. It is interesting and a topic of current research, you are encouraged to read it and also check out the discussions of the electron and muon $g - 2$ measurements at http://www.astro.caltech.edu/~golwala/ph135c/.

But, let's continue onward and discuss the implications of this Hamiltonian. For now, we will take $H = \mathbf{H}_{int}$ and neglect the orbital motion. Since spin is not connected to spatial momentum or orbital angular momentum, we guess that it will be safe to assume they decouple, but we will justify this later. For now, you may think of this as a toy problem, or consider the particle with the spin to be constrained so it may rotate but not change its spatial position. Under this assumption, let's calculate the unitary time evolution operator:

$$\mathbf{U}(t) = \exp\left(-\frac{i}{\hbar}\,\mathbf{H}\,t\right) = \exp\left(\frac{i}{\hbar}\,\gamma\,t\,\vec{\mathbf{S}}\cdot\vec{B}\right) \tag{14.116}$$

Since $\vec{B}$ is constant, we see that our time-evolution operator is the same as the one that would actively rotate the state by an angle $-\gamma\,B\,t = -\omega_0\,t$ about $\vec{B}$.

Let's explicitly calculate **U**(t) for the case $\vec{B} = B\,\widehat{z}$:

$$\mathbf{U}(t) = \exp\left(\frac{i}{\hbar}\,\gamma\,t\,\vec{\mathbf{S}} \cdot B\widehat{z}\right) \tag{14.117}$$

$$= \mathbf{I}\cos\left(-\frac{\omega_0\,t}{2}\right) - \boldsymbol{\sigma}_z\,i\,\sin\left(-\frac{\omega_0\,t}{2}\right) \tag{14.118}$$

$$= \left[\begin{array}{cc} e^{\,i\,\omega_0 t/2} & 0 \\ 0 & e^{-i\,\omega_0 t/2} \end{array}\right] \tag{14.119}$$

Let's apply it to the state $|\widehat{n}, + \rangle$, with $\widehat{n}$ not necessarily along $\widehat{z}$:

$$\mathbf{U}(t)|\widehat{n}, + \rangle = \left[\begin{array}{cc} e^{\,i\,\omega_0 t/2} & 0 \\ 0 & e^{-i\,\omega_0 t/2} \end{array}\right] \left[\begin{array}{c} e^{-i\,\phi/2}\cos\frac{\theta}{2} \\ e^{\,i\,\phi/2}\sin\frac{\theta}{2} \end{array}\right] \tag{14.120}$$

$$= \left[\begin{array}{c} e^{-i(\phi-\omega_0 t)/2}\cos\frac{\theta}{2} \\ e^{\,i(\phi-\omega_0 t)/2}\sin\frac{\theta}{2} \end{array}\right] \tag{14.121}$$

which, clearly, corresponds to CW rotation about $\widehat{z}$ by $\omega_0\,t$. That is, we get precession of the quantum mechanical spin as expected.

The precession is clear if we calculate $\langle \vec{\mathbf{S}} \rangle$ as a function of time using the above:

$$\langle \mathbf{S}_z \rangle = \left[\begin{array}{cc} e^{i(\phi-\omega_0 t)/2} \cos \frac{\theta}{2} & e^{-i(\phi-\omega_0 t)/2} \sin \frac{\theta}{2} \end{array}\right] \frac{\hbar}{2} \left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right] \left[\begin{array}{c} e^{-i(\phi-\omega_0 t)/2} \cos \frac{\theta}{2} \\ e^{i(\phi-\omega_0 t)/2} \sin \frac{\theta}{2} \end{array}\right]$$
$$(14.122)$$

$$= \frac{\hbar}{2} \left(\cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2} \right) = \frac{\hbar}{2} \cos \theta \qquad (14.123)$$

$$\langle \mathbf{S}_x \rangle = \left[\begin{array}{cc} e^{i(\phi-\omega_0 t)/2} \cos \frac{\theta}{2} & e^{-i(\phi-\omega_0 t)/2} \sin \frac{\theta}{2} \end{array}\right] \frac{\hbar}{2} \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right] \left[\begin{array}{c} e^{-i(\phi-\omega_0 t)/2} \cos \frac{\theta}{2} \\ e^{i(\phi-\omega_0 t)/2} \sin \frac{\theta}{2} \end{array}\right]$$
$$(14.124)$$

$$= \frac{\hbar}{2} \cos \frac{\theta}{2} \sin \frac{\theta}{2} \left(e^{i(\phi-\omega_0 t)} + e^{-i(\phi-\omega_0 t)} \right) = \frac{\hbar}{2} \sin \theta \cos (\phi - \omega_0 t) \quad (14.125)$$

$$\langle \mathbf{S}_y \rangle = \left[\begin{array}{cc} e^{i(\phi-\omega_0 t)/2} \cos \frac{\theta}{2} & e^{-i(\phi-\omega_0 t)/2} \sin \frac{\theta}{2} \end{array}\right] \frac{\hbar}{2} \left[\begin{array}{cc} 0 & -i \\ i & 0 \end{array}\right] \left[\begin{array}{c} e^{-i(\phi-\omega_0 t)/2} \cos \frac{\theta}{2} \\ e^{i(\phi-\omega_0 t)/2} \sin \frac{\theta}{2} \end{array}\right]$$
$$(14.126)$$

$$= -i \frac{\hbar}{2} \cos \theta \sin \theta \left(e^{i(\phi-\omega_0 t)} + e^{-i(\phi-\omega_0 t)} \right) = \frac{\hbar}{2} \sin \theta \sin (\phi - \omega_0 t) \quad (14.127)$$

That is, the expectation value is a vector of length $\hbar/2$ tilted at polar angle θ and that precesses in ϕ CW about $\hat{z}$ at angular speed $\omega_0 t$.

Lecture 39:
Applications of Spin Angular Momentum

Revision Date: 2008/02/04

Applications of Spin Angular Momentum

Spin Hamiltonian Only: Paramagnetic Resonance

Our first application of spin angular momentum will be paramagnetic resonance, in which we still only need to consider particle states in $\tau^{(1/2)}$, but for which the dynamics is richer than simple spin precession.

Let us first set up the classical problem. Consider a magnetic moment $\vec{\mu}$ in a uniform magnetic field $\vec{B}_0 = B_0 \, \widehat{z}$. We do not yet need to specify the magnitude of the spin. As we proved above, the spin precesses about the $\vec{B}_0$ direction at angular frequency $\vec{\omega}_0 = -\gamma \, \vec{B}_0$.

Let's work in a frame F_r that is rotating at frequency $\vec{\omega}$. We shall designate the inertial lab frame by F. (Since we have no spatial degrees of freedom, the change to a noninertial frame does not result in any fictitious forces or potentials that we need to add to the Hamiltonian.) In this frame, the spin appears to precess with angular velocity

$$\vec{\omega}_r = \vec{\omega}_0 - \vec{\omega} = -\gamma \left(\vec{B}_0 + \frac{\vec{\omega}}{\gamma} \right) \tag{14.128}$$

Thus, in F_r, the effective magnetic field is

$$\vec{B}_r = \vec{B}_0 + \frac{\vec{\omega}}{\gamma} \tag{14.129}$$

We say "effective" because, given that we are *only* considering the dynamics of the spin, there is no way to measure $\vec{B}$ other than to measure the spin precession frequency.

Applications of Spin Angular Momentum (cont.)

Now, consider application of an additional magnetic field $\vec{B}(t)$ that is in a plane perpendicular to $\vec{B}_0$, that is rotating CW in that plane with angular frequency ω (and hence with angular frequency vector $\vec{\omega} = -\omega\,\hat{z}$), and that is small in magnitude compare to B_0. In the inertial lab frame F, the total magnetic field is

$$\vec{B}(t) = \hat{x}\,B\,\cos\omega\,t - \hat{y}\,B\,\sin\omega\,t + B_0\,\hat{z} \qquad \text{with} \qquad B \ll B_0 \qquad (14.130)$$

Let's take the initial condition for the classical problem to be $\vec{\mu}$ initially pointed along $\vec{B}_0$, $\vec{\mu}(t=0) = \mu\,\hat{z}$.

Let's determine the classical evolution of the magnetic moment in a rotating frame F_r that rotates with the same angular frequency vector $\vec{\omega}$ as the rotating field. The component of the field that was rotating in the lab frame will clearly appear fixed. The fixed component of the field $\vec{B}_0 = B_0\,\hat{z}$ in the lab frame will be reduced in effective size because of the precession correction given earlier. *We are using here the approximation $B \ll B_0$ so that we may ignore the rotating field in calculating the precession due to $\vec{B}_0$.* We obtain that the effective field in the rotating frame is

$$\vec{B}_r = B\,\hat{x}_r + \left(B_0 - \frac{\omega}{\gamma}\right)\hat{z} \qquad (14.131)$$

where $\hat{x}_r$ is the unit vector in the x direction of the *rotating* frame. Note that $\hat{z}_r = \hat{z}$ because the rotation is about $\hat{z}$.

Now, we see that, because the effective field direction is no longer along $\widehat{z} = \widehat{z}_r$, the precession will no longer be about $\widehat{z}$ and the precession frequency will have changed. Rather, the spin will precess about $\vec{B}_r$, which is in the $x_r z_r$ plane and makes a polar angle

$$\cos \alpha = \frac{\left(B_0 - \frac{\omega}{\gamma}\right)}{\left[B^2 + \left(B_0 - \frac{\omega}{\gamma}\right)^2\right]} \tag{14.132}$$

with $\widehat{z}_r = \widehat{z}$. Since the spin is initially along $\widehat{z} = \widehat{z}_r$, its tip follows a cone that must intersect the $\widehat{z}$ axis at $t = 0$, so the precession cone has opening angle α and $\vec{\mu}$ always makes this angle with $\vec{B}_r$. The precession frequency will be given by the standard formula, $\vec{\omega}_r = -\gamma \vec{B}_r$, but, because $B_r \neq B_0$, the precession frequency's magnitude is now

$$\omega_r = \gamma B_r = \gamma \sqrt{B^2 + \left(B_0 - \frac{\omega}{\gamma}\right)^2} \tag{14.133}$$

We may calculate explicitly the evolution $\vec{\mu}(t)$ by recognizing that, because $\vec{B}_r$ is fixed and along the z-axis of a frame that is tilted from F_r by a polar angle α, $\vec{\mu}$ does simple precession about the z axis in that frame. So we just need to write down the time evolution of standard precession and combine it with rotation back to the F_r frame by an angle α about $-\widehat{y}_r$ (Note the negative sign on $\widehat{y}_r$!). The result is:

$$\left(\vec{\mu}(t)\right)_r = \mu \left[\begin{array}{ccc} c_\alpha & 0 & -s_\alpha \\ 0 & 1 & 0 \\ s_\alpha & 0 & c_\alpha \end{array} \right] \left[\begin{array}{c} s_\alpha \cos \omega_r t \\ s_\alpha \sin \omega_r t \\ c_\alpha \end{array} \right] = \mu \left[\begin{array}{c} c_\alpha \, s_\alpha \, (\cos \omega_r t - 1) \\ s_\alpha \, \sin \omega_r t \\ c_\alpha^2 + s_\alpha^2 \cos \omega_r t \end{array} \right]$$

$$(14.134)$$

where we write $\vec{\mu}(t)$ with an underline and $_r$ subscript because we are calculating its coordinate representation in F_r. The column vector in the first expression is the coordinate representation for $\vec{\mu}(t)$ in a frame whose z axis is $\vec{B}_r$, expressing the simple precession about $\vec{B}_r$ ($c_\alpha = \cos \alpha$ and $s_\alpha = \sin \alpha$ as usual). The rotation matrix transforms the coordinate representation from that frame to F_r, which is done by a rotation by α CW about $\widehat{y}_r$. It is easy to see that the final result is consistent with $\vec{\mu}(t = 0) = \mu \, \widehat{z}_r$ as expected. However, $\vec{\mu} \cdot \widehat{z}_r$ now oscillates with time at frequency ω_r.

Applications of Spin Angular Momentum (cont.)

The behavior in the lab frame is easy to calculate explicitly by applying the appropriate rotation matrix to recover the coordinate representation in F. That matrix is specified by the fact that F_r rotates about $\hat{z}$ at $-\omega$. That is,

$$(\vec{\mu}(t)) = \mu \begin{bmatrix} \cos \omega t & \sin \omega t & 0 \\ -\sin \omega t & \cos \omega t & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_\alpha \, s_\alpha \, (\cos \omega_r t - 1) \\ s_\alpha \, \sin \omega_r t \\ c_\alpha^2 + s_\alpha^2 \cos \omega_r t \end{bmatrix} \tag{14.135}$$

One can check the sign of the rotation matrix by checking that applying it to F_r's x axis, $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$, which is the direction of the rotating field in F_r, returns $\begin{bmatrix} \cos \omega t & -\sin \omega t & 0 \end{bmatrix}^T$, which is the direction of the rotating field in F by our initial assumption. The result is

$$(\vec{\mu}(t)) = \mu \begin{bmatrix} s_\alpha \, [c_\alpha \cos \omega t \, (\cos \omega_r t - 1) + \sin \omega t \sin \omega_r t] \\ s_\alpha \, [c_\alpha \sin \omega t \, (1 - \cos \omega_r t) + \cos \omega t \cos \omega_r t] \\ c_\alpha^2 + s_\alpha^2 \cos \omega_r t \end{bmatrix} \tag{14.136}$$

Note that the z component of $\vec{\mu}(t)$ in F and F_r are the same, which makes sense since the two are related by rotation about z. We could have guessed this, but it was instructive to transform the entire vector for practice.

In looking at the z component of the above, which is just $\vec{\mu}(t) \cdot \hat{z}$, we see that can get full modulation of $\vec{\mu}(t) \cdot \hat{z}$ — that is, make its value oscillate between μ and $-\mu$ — if $\alpha = \pi/2$. The condition for that to happen is $\cos \alpha = 0$, which occurs when $\omega = \gamma B_0$. What is happening physically when this occurs?

▶ By setting $\omega = \gamma B_0$, we are requiring that the applied rotating field rotate at the same frequency about $\vec{B}_0$ as the moment $\vec{\mu}$ would precess about $\vec{B}$ in the absence of a rotating field. That is, the rotating field is synchronized with the nominal dominant precession of $\vec{\mu}$.

▶ Therefore, in F_r, the effective field $\vec{B}_r$ is entirely along $\hat{x}_r$; the z component vanishes because of the choice of ω. Because the z component of $\vec{B}_r$ vanishes, $\vec{\mu}$ does not precess about z in this frame, as expected from the synchronization pointed out above.

▶ But, now, in F_r, $\vec{\mu}$ sees $\vec{B} = B \hat{x}_r$ as a time-independent uniform field. Therefore, it precesses about $\hat{x}_r$ with precession frequency ω_r, which, for $\omega = \gamma B_0$, is $\omega_r = \gamma B$. That is, we get standard precession of $\vec{\mu}$ about $\vec{B} = B \hat{x}_r$ in F_r. Since $\alpha = \pi/2$ to obtain this condition, $\vec{\mu}$ precesses in the $y_r z_r$ plane, making an angle $\pi/2$ with $\hat{x}_r$.

Applications of Spin Angular Momentum (cont.)

▶ Now, consider applying the rotating field only for a finite time $\tau = (\pi/2)/\omega_r$ — *i.e.*, 1/4 of the period for precessing about $\vec{B} = B\,\widehat{x}_r$. This is called a $\pi/2$ *pulse*. Then $\vec{\mu}$ will precess from begin parallel to $\widehat{z}_r$ to being parallel to $\widehat{y}_r$. When the rotating field is turned off, the moment will lie in the *xy* plane and will now see only the time-independent field $\vec{B}_0 = B_0\,\widehat{z}$. It will therefore now precess about $\vec{B}_0$ with the standard precession frequency $\omega_0 = \gamma\,B_0$.

▶ If, instead, we applied the rotating field for $\tau = \pi/\omega_r$ — *i.e.*, 1/2 of the period for precessing about $\vec{B} = B\,\widehat{x}_r$, also known as a π *pulse* — then the moment will precess from parallel to $\widehat{z}_r$ to parallel to $-\widehat{z}_r$. When the rotating field is turned off, the spin will be flipped from its initial condition parallel to $\widehat{z}$ to antiparallel to $\widehat{z}$.

The QM version of this for $\omega = \omega_0$ will be assigned as a homework problem. You will see that the expectation value of **S** follows the above behavior. Moreover, you will see that a $\pi/2$ pulse evolves the moment from the state $|\widehat{z}, + \rangle$ to the state $|\widehat{n}, + \rangle$ where $\widehat{n}$ is perpendicular to $\widehat{z}$ (how it decomposes in terms of $\widehat{x}$ and $\widehat{y}$ depends on ω_0, which is the speed at which the pulsed field rotates for resonance to occur, while ω_r, the frequency that determins the length of the $\pi/2$ pulse, is determined by the magnitude of the pulsed field *B*.), and that a π pulse evolves the state from $|\widehat{z}, + \rangle$ to $|\widehat{z}, - \rangle$.

The term paramagnetic resonance is used because, with these $\pi/2$ or π pulses, one can move the system from one eigenstate fully into another one — the probability of obtaining a particular value of $\widehat{n} \cdot \vec{\mathbf{S}}$ for specific $\widehat{n}$ becomes 100%.

Uses of paramagnetic resonance:

▶ Nuclear magnetic resonance/magnetic resonance Imaging
Since this is such well-studied technique, you can find lots of information on this on the web. See, for example, the writeup from the Ph77 labs, http://www.pma.caltech.edu/˜ph77/labs/nmr.pdf.

▶ Spin dynamics in nEDM experiment
One modern application of paramagnetic resonance is in an experiment to search for an *electric* dipole moment of the neutron (nEDM). The only know vector quantity associated with a subatomic particle like the neutron is its spin. Since a nEDM is a vector, is must be proportional to the spin and hence to the neutron magnetic moment, $\vec{d}_n \propto \vec{\mu}_n$. The energy of an electric dipole moment in an electric field is the same as that of a magnetic moment with a magnetic field, $U = -\vec{d} \cdot \vec{E}$. This interaction is odd under time-reversal – $\vec{E}$ does not change sign (it is just determined by the location of charges) but $\vec{d}$ does if it is set by the spin because spin changes sign under time reversal. (Just think about how an angular frequency changes sign under time reversal.) Such interactions are highly suppressed in the standard model of particle physics, but are seen to occur at a very low level. Models of physics beyond the standard model tend to predict larger amounts of *T*-violation. A larger amount of *T*-violation than has been observed is a requisite for explaining the matter-antimatter asymmetry in the universe. Expectations are for a nEDM at the level of 10^{-27} *e*-cm. Searches to date have reached to $\sim 10^{-25}$ *e*-cm, perhaps a factor of two lower. Deeper searches for a nEDM are thus very well motivated by particle physics.

A new technique being developed to do such a search makes use of the fact that, just as a magnetic field causes spin precession, an electric field acting on a nEDM will cause spin precession because the nEDM is tied to the particle spin. That is, the electric field causes a torque

$$\vec{T} = \vec{d} \times \vec{E} \tag{14.137}$$

to make the EDM want to align with the $\vec{E}$ field. Of course, if the EDM is tied to a spin, the spin appears on both the left and right sides of the equation and one gets precession instead of simple rotational motion about an axis parallel to $\vec{d}$ and $\vec{E}$.

The way the experiment is done is to confine a sample of ultra-cold neutrons (UCN, $E < 0.13\mu$eV, $v < 5$ m/s) in a region of parallel magnetic and electric field. The natural spin precession frequency, ω_0 in the above, is now, with the electric field

$$\omega_{0,\pm} = -\left(\frac{2}{\hbar}\mu_n B_0 \pm \frac{2}{\hbar} d_n E_0\right) \tag{14.138}$$

where the $\pm$ sign depends on whether $\vec{E}_0$ is parallel or antiparallel to $\vec{B}_0$. The methodology is clear, then: monitor the spin precession frequency and look for a difference in frequency between the two cases of $\vec{E}_0$ parallel and antiparallel to $\vec{B}_0$. The spin precession frequency is obtained by monitoring the the AC magnetic field emanating from the neutron sample: because $\vec{\mu}$ is precessing, the magnetization vector (magnetic moment) of the sample precesses, and that varying magnetic field can be monitored with a pickup coil (usual read out with a SQUID, a very sensitive magnetometer).

The UCN spins must of course be aligned in a very specific manner for the experiment to work. The UCN first brought in with their spins parallel to the static $\vec{B}$ field. This puts all the UCN in the same initial state, with $\vec{\mu}$ (and therefore $\vec{d}$ along $\hat{z}$, the field direction). Then, a $\pi/2$ pulse is applied to rotate the UCN spin vectors to point in the $\hat{y}_r$ direction. The UCN spins are now in the xy plane relative to the field, so they precess at ω_0 given above. Two adjacent cells with opposite $\vec{E}$ and the same $\vec{B}$ direction are observed. The UCN are lost on the timescale of about 10 min (free decay, capture, losses at the walls). The UCN cell is refilled and the measurement repeated over and over. The $\vec{B}$ and $\vec{E}$ field directions are switched to check for systematica errors. The $\vec{B}$ field is monitored by simultaneous observation of ^{3}He spin precession in the same cells; ^{3}He has no EDM, so it monitors $\vec{B}$ for systematics that might introduce a fake nEDM signal. It requires a bit of work to find a RF field ω for which a $\pi/2$ pulse works for the different μ of the neutrons and ^{3}He, but it is possible.

More details can be found at http://p25ext.lanl.gov/edm/edm.html.

Applications of Spin Angular Momentum (cont.)

Spatial + Spin Hamiltonian: Zeeman Effect in the SHO

Let's now consider a case in which the Hamiltonian includes operators in both that scalar Hilbert space $\mathcal{V}$ as well as in the spin space $\tau^{(1/2)}$, but for which the two terms are separable. That is,

$$H = H_0 \otimes \mathbf{I} + I \otimes H_s \tag{14.139}$$

where H_0 acts in the scalar Hilbert space, H_s acts in $\tau^{(1/2)}$, $\mathbf{I}$ is the identity operator in $\tau^{(1/2)}$, and I is the identity operator is the scalar Hilbert space $\mathcal{V}$. The two terms in the Hamiltonian commute (because each one is the identity in the space in which the other one is not), so eigenstates of the overall Hamiltonian are simultaneous eigenstates of the two pieces.

Let's begin by letting H_0 be the Hamiltonian of the three-dimensional SHO. This problem is worked out in detail in Shankar 12.6. For our purposes, it suffices to know that, in spherical coordinates, the Hamiltonian has eigenstates in the scalar Hilbert space $\mathcal{V}$ labeled by three quantum numbers $|n, l, m\rangle$, with

$$E_{n,l,m} = \left(n + \frac{3}{2}\right)\hbar\omega \qquad l = n, n-2, \ldots, 1 \text{ or } 0 \tag{14.140}$$

$$\langle \vec{r} | n, l, m \rangle = \psi_{n,l,m}(r, \theta, \phi) = R_{E,l}(r)\, Y_l^m(\theta, \phi) \tag{14.141}$$

There is a good deal of degeneracy in that multiple values of l give the same energy.

Let us initially assume that $H_s = 0$ — that is, the energy does not depend on the spin state — so that the product-space Hamiltonian is

$$H = H_{SHO} \otimes I \qquad (14.142)$$

where H_{SHO} is the SHO Hamiltonian acting in $\mathcal{V}$ and I is the identity operator in $\tau^{(1/2)}$. The states are in the product space $V = \mathcal{V} \otimes \tau^{(1/2)}$. The choice of basis in $\tau^{(1/2)}$ is arbitrary since the Hamiltonian is the identity operator in $\tau^{(1/2)}$ (every state in $\tau^{(1/2)}$ is an eigenstate of the piece of H that acts in $\tau^{(1/2)}$, I.) Let's arbitrarily take the eigenstates of $\mathbf{S}_z$ as the basis for $\tau^{(1/2)}$, so the eigenstates of the product-space H are

$$|n, l, m, m_s\rangle = |n, l, m\rangle \otimes |m_s\rangle \qquad (14.143)$$

where $m_s = \pm 1/2$, $|\pm 1/2\rangle = |\hat{z}, \pm\rangle$. The energies and allowed values of l are unchanged because H has no piece acting in $\tau^{(1/2)}$.

Applications of Spin Angular Momentum (cont.)

Now, let's add an interaction with a magnetic field $\vec{B} = B\,\widehat{z}$. The Hamiltonian becomes

$$H = H_{SHO} \otimes I + \left(-\vec{\mu}_l \cdot \vec{B} \right) \otimes I + I \otimes \left(-\vec{\mu}_s \cdot \vec{B} \right) \tag{14.144}$$

$$= H_{SHO} \otimes I + \left(-\frac{q\,B}{2\,M\,c}\,L_z \right) \otimes I + I \otimes \left(-\frac{q\,B}{M\,c}\,\mathbf{S}_z \right) \tag{14.145}$$

where $\vec{\mu}_l$ and $\vec{\mu}_s$ are the magnetic moments due to orbital and spin angular momentum, and M is the particle mass. The additional terms that enter because of the magnetic field commute with the original H_{SHO} term: the L_z term acts in $\mathcal{V}$ but commutes with H_{SHO} because H_{SHO} is spherically symmetric; and $\mathbf{S}_z$ acts in $\tau^{(1/2)}$, not $\mathcal{V}$, so is the identity operator in the latter. Therefore, the eigenstates in $\mathcal{V}$ we chose for the $B = 0$ case remain valid, and, the basis we chose for $\tau^{(1/2)}$ before was a good choice because they are now eigenstates of H. However, the energies change. We have

$$E_{n,l,m,m_s} = \left(n + \frac{3}{2} \right) \hbar\omega - \frac{q\,B}{2\,M\,c}\,m\,\hbar - \frac{q\,B}{M\,c}\,m_s\,\hbar \tag{14.146}$$

There remains a good deal of degeneracy, but states with the same n acquire different energies if their values of $m + 2\,m_s$ are different. Certainly, states with the same n, l, m but different m_s have different energies and states with the same n, l, m_s but different m are also split – these splittings are what we term the Zeeman effect. As is explained in Shankar 14.5, a similar phenomenon occurs in the hydrogen atom.

As Shankar also indicates, another type of Hamiltonian that combines spatial and spin effects is the one for spin-orbit coupling,

$$H_{\vec{L}\cdot\vec{S}} = a\vec{L}\cdot\vec{S} = a\sum_i L_i \otimes \mathbf{S}_i \qquad (14.147)$$

Physically, this term arises by considering the magnetic field seen by the electron's magnetic moment in its rest frame due to the proton's apparent motion in that frame. The interesting thing here is that we see that this term has factors in both $\mathcal{V}$ and $\tau^{(1/2)}$, and, moreover, not all the terms commute with $L_z \otimes \mathbf{I}$ and $I \otimes \mathbf{S}_z$. We will have to deal with it using perturbation theory rather than by finding exact eigenstates of the total Hamiltonian that includes the spatial term (hydrogen atom or 3D SHO, for example) and this term.

Stern-Gerlach Experiment

See Shankar – this is sufficiently straightforward that repeating it here is unnecessary.

Section 15
Addition of Angular Momenta

Lecture 40:
Addition of Angular Momentum – States

Revision Date: 2008/03/02

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:

$$V^{(j_1)} \otimes V^{(j_2)} = V^{(j_1+j_2)} \oplus V^{(j_1+j_2-1)} \oplus \cdots \oplus V^{(j_1-j_2+1)} \oplus V^{(j_1-j_2)} \qquad (15.1)$$

- 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 hint of this in our demonstration in Section 14.2 that, when we try to add "spin" information to a particle state by considering direct products of a scalar Hilbert space $\mathcal{V}$ and a classical vector or spin-1/2 spinor space $\tau^{(j)}$ ($j = 1$ for a vector, $j = 1/2$ for a spin-1/2 spinor), we find that the generator of rotations for the product space is

$$\vec{\mathcal{J}} = \vec{L} \otimes I^{(j)} + I \otimes \vec{\mathbf{J}}^{(j)} \tag{15.2}$$

$$\vec{\mathbf{J}}^{(1)} = i\,\hbar\,\vec{\mathbf{M}} \qquad \vec{\mathbf{J}}^{(1/2)} = \vec{\mathbf{S}} = \frac{\hbar}{2}\vec{\boldsymbol{\sigma}} \tag{15.3}$$

We have so far written states in this product space in terms of the eigenstates of L^2, $\vec{L}$, $(\vec{\mathbf{J}}^{(j)})^2$, and $\mathbf{J}_z^{(j)}$. 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 $(\vec{\mathcal{J}})^2$ 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 $\mathcal{V}^{(1)}$ and $\mathcal{V}^{(2)}$ (the superscript now indicates which particle, not a j value), then the generator of rotations in the product space is

$$\vec{\mathcal{J}} = \vec{L}^{(1)} \otimes I^{(2)} + I^{(1)} \otimes \vec{L}^{(2)} \tag{15.4}$$

Another example would be to consider the combined spin of a two-particle system consisting of spins j_1 and j_2:

$$\vec{\mathcal{J}} = \vec{\mathbf{J}}^{(j_1)} \otimes \mathbf{I}^{(j_2)} + \mathbf{I}^{(j_1)} \otimes \vec{\mathbf{J}}^{(j_2)} \tag{15.5}$$

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 13.6, 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 the notational complications (boldface, calligraphic letters) because we do not care about the distinction between orbital, spin, and total angular momentum right now. So we consider

$$\vec{J} = \vec{J}^{(j_1)} \otimes I^{(2)} + I^{(1)} \otimes \vec{J}^{(j_2)} \qquad V = V^{(j_1)} \otimes V^{(j_2)} \tag{15.6}$$

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,

$$[J_a, J_b] = i\,\hbar\,\epsilon_{abc} J_c \tag{15.7}$$

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 $m = j, j-1, \ldots, -j+1, -j$; and 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 $\{|j_1, m_1\rangle \otimes |j_2, m_2\rangle\}$.

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 $V^{(j_1+j_2)}$; that is, $V^{(j_1+j_2)}$ is a subspace of V. Moreover, just by counting states, we know that $V^{(j_1+j_2)}$ cannot be all of V: V has $(2j_1 + 1)(2j_2 + 1)$ states, while $V^{(j_1+j_2)}$ 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 V. Moreover, because of the angular momentum structure of $\vec{J}$, those subspaces must be $V^{(j)}$ for some values of j to be determined. That is, we already know

$$V^{(j_1)} \otimes V^{(j_2)} = V^{(j_1+j_2)} \oplus \sum_{j \ tbd} V^{(j)} \tag{15.8}$$

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 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

$$
\begin{array}{ll}
J_z & \text{contributing } (m_1, m_2) \text{ values} \\
j_1 + j_2 & (j_1, j_2) \\
j_1 + j_2 - 1 & (j_1, j_2 - 1), \ (j_1 - 1, j_2) \\
j_1 + j_2 - 2 & (j_1, j_2 - 2), \ (j_1 - 1, j_2 - 1), \ (j_1 - 2, j_2) \\
\quad \vdots & \vdots \\
j_1 + j_2 - n & (j_1, j_2 - n), \ (j_1 - 1, j_2 - (n-1)), \ldots, (j_1 - (n-1), j_2 - 1), \ (j_1 - n, j_2) \\
\quad \vdots & \vdots \\
j_1 - j_2 & (j_1, -j_2), \ (j_1 - 1, -j_2 + 1), \ldots, (j_1 - 2j_2 + 1, j_2 - 1), \ (j_1 - 2j_2, j_2) \\
j_1 - j_2 - 1 & (j_1 - 1, -j_2), \ (j_1 - 2, -j_2 + 1), \ldots, (j_1 - 1 - 2j_2 + 1, j_2 - 1), \ (j_1 - 1 - 2j_2, j_2) \\
\quad \vdots & \vdots
\end{array}
$$

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 $n = 0, \ldots, 2j_2$. $V^{(j_1 + j_2)}$ can only yield 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 $V^{(j)}$ for $j \geq j_1 + j_2 - 1$ can yield such a state. But if $j \geq j_1 + j_2$, then $V^{(j)}$ 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 $V^{(j_1 + j_2 - 1)}$. A similar argument holds for the remaining $n = 2, 3, \ldots 2j_2$. So we successively add $V^{(j_1 + j_2 - 2)}$, $V^{(j_1 + j_2 - 3)}$, ..., $V^{(j_1 + j_2 - 2j_2)} = V^{(j_1 - j_2)}$. That is,

$$V^{(j_1)} \otimes V^{(j_2)} = V^{(j_1 + j_2)} \oplus V^{(j_1 + j_2 - 1)} \oplus \cdots \oplus V^{(j_1 - j_2 + 1)} \oplus V^{(j_1 - j_2)} \oplus \cdots \quad (15.9)$$

Addition of Angular Momentum – States (cont.)

▶ Moreover, we can conclude that the direct sum terminates at $V^{(j_1 - j_2)}$, that the $\cdots$ at the end are unnecessary, by just counting states. The space $V^{(j)}$ contributes $2j + 1$ states. The total number of states we have accounted for so far is

$$N = \sum_{j=j_1-j_2}^{j_1+j_2} 2j + 1 = \sum_{j=0}^{j_1+j_2} (2j+1) - \sum_{j=0}^{j_1-j_2-1} (2j+1) \tag{15.10}$$

$$= (j_1 + j_2)(j_1 + j_2 + 1) + (j_1 + j_2 + 1) \tag{15.11}$$
$$\quad - (j_1 - j_2 - 1)(j_1 - j_2) - (j_1 - j_2)$$

$$= 4 j_1 j_2 + 2 j_1 + 2 j_2 + 1 = (2 j_1 + 1)(2 j_2 + 1) \tag{15.12}$$

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 $2 j_2 + 1$ terms in the direct sum,

$$V^{(j_1)} \otimes V^{(j_2)} = V^{(j_1+j_2)} \oplus V^{(j_1+j_2-1)} \oplus \cdots \oplus V^{(j_1-j_2+1)} \oplus V^{(j_1-j_2)} \tag{15.13}$$

▶ For completeness, 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 $m < j_1 - j_2$, 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

$$E_{n,l,m} = \left(n + \frac{3}{2}\right)\hbar\omega \qquad l = n, n-2, \ldots, 1 \text{ or } 0 \tag{15.14}$$

So l can take on integer values. Consider the $n = 1$ state, which allows $l = 1$ 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 $j = 1/2$.

▶ 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 $V^{(j)}$ where it will cause no confusion.
  We have

$$\frac{1}{2} \otimes \frac{1}{2} \otimes \frac{1}{2} = \frac{1}{2} \otimes (1 \oplus 0) = \left(\frac{1}{2} \otimes 1\right) \oplus \left(\frac{1}{2} \otimes 0\right) \tag{15.15}$$

$$= \left(\frac{3}{2} \oplus \frac{1}{2}\right) \oplus \frac{1}{2} = \frac{3}{2} \oplus \frac{1}{2} \oplus \frac{1}{2} \tag{15.16}$$

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 $V^{(1/2)}$. 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 V.

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 $l = 0$ will change nothing, adding $l = 1$ 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 $l = 0$, 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 (because the quark content of thes particles us 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 Paul exclusion restriction.

States with $l \neq 0$ may manifest as stable particles, or simply as short-lived resonances.

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,

$$|j, m\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} \left(|j_1, m_1\rangle \otimes |j_2, m_2\rangle\right) \left(\langle j_1, m_1 | \otimes \langle j_2, m_2 |\right) |j, m\rangle \qquad (15.17)$$

where the expansion coefficients are called the Clebsch-Gordan (CG) coefficients.

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.

In detail:

1. First, we know only one uncoupled basis state corresponds to the coupled basis state of maximum m:

$$|j = j_1 + j_2, m = j_1 + j_2 \rangle = |j_1, j_1 \rangle \otimes |j_2, j_2 \rangle \tag{15.18}$$

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_-:

$$
\begin{aligned}
&J_- |j = j_1 + j_2, m = j_1 + j_2\rangle \\
&= \left(J_-^{(j_1)} \otimes I^{(j_2)} + I^{(j_1)} \otimes J_-^{(j_2)} \right) \left(|j_1, j_1\rangle \otimes |j_2, j_2\rangle \right) \\
&= \hbar \sqrt{(j_1 + j_1)(j_1 - j_1 + 1)} \, |j_1, j_1 - 1\rangle \otimes |j_2, j_2\rangle \\
&\quad + \hbar \sqrt{(j_2 + j_2)(j_2 - j_2 + 1)} \, |j_1, j_1\rangle \otimes |j_2, j_2 - 1\rangle \\
&= \hbar \sqrt{2 j_1} \, |j_1, j_1 - 1\rangle \otimes |j_2, j_2\rangle + \hbar \sqrt{2 j_2} \, |j_1, j_1\rangle \otimes |j_2, j_2 - 1\rangle
\end{aligned}
$$

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

$$
|j = j_1 + j_2, m = j_1 + j_2 - 1\rangle = \frac{\sqrt{j_2} \, |j_1, j_1\rangle \otimes |j_2, j_2 - 1\rangle + \sqrt{j_1} \, |j_1, j_1\rangle \otimes |j_1 - 1, j_2\rangle}{\sqrt{j_1 + j_2}}
$$

Continuing downward is rather tedious but otherwise straightforward. $\qquad$ (15.19)

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/).

Lecture 41:
Addition of Angular Momentum – States

Revision Date: 2008/02/08

Addition of Angular Momentum – States

General Properties of Clebsch-Gordan Coefficients

▶ A CG coefficient is nonvanishing only if $j_1 - j_2 \leq j \leq j_1 + j_2$

$$(\langle j_1, m_1 \mid \otimes \langle j_2, m_2 \mid) \neq 0 \quad \text{only for} \quad j_1 - j_2 \leq j \leq j_1 + j_2 \tag{15.20}$$

▶ A CG coefficient is nonvanishing only if $m = m_1 + m_2$.

▶ By convention, CG coefficients are always real.

▶ By convention,

$$(\langle j_1, j_1 \mid \otimes \langle j_2, j - j_1 \mid) \mid j, j \rangle \geq 0 \quad \text{for any} \quad j_1 - j_2 \leq j \leq j_1 + j_2 \tag{15.21}$$

This fixes the sign of the top state for each j.

▶ The symmetry properties under $m_1, m_2, m \rightarrow -m_1, -m_2, -m$ imply

$$(\langle j_1, -m_1 \mid \otimes \langle j_2, -m_2 \mid) \mid j, -m \rangle = (-1)^{j_1 + j_2 - j} (\langle j_1, m_1 \mid \otimes \langle j_2, m_2 \mid) \mid j, m \rangle \tag{15.22}$$

▶ 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

$$|j = 1, m = 1 \rangle = \left| \frac{1}{2}, \frac{1}{2} \right\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle \tag{15.23}$$

$$|j = 1, m = -1 \rangle = \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \tag{15.24}$$

We use the formula given above for the action of the lowering operator on the top state to obtain

$$|j = 1, m = 0 \rangle = \frac{1}{\sqrt{2}} \left(\left| \frac{1}{2}, \frac{1}{2} \right\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle + \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle \right) \tag{15.25}$$

Finally, we use the orthogonality requirement and the normalization, realness, sign-fixing conventions to obtain

$$|j = 0, m = 0 \rangle = \frac{1}{\sqrt{2}} \left(\left| \frac{1}{2}, \frac{1}{2} \right\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle - \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle \right) \tag{15.26}$$

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:

$$\left| \frac{1}{2}, \frac{1}{2} \right\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = \frac{1}{\sqrt{2}} \left(|j = 1, m = 0 \rangle + |j = 0, m = 0 \rangle \right) \tag{15.27}$$

$$\left| \frac{1}{2}, -\frac{1}{2} \right\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle = \frac{1}{\sqrt{2}} \left(|j = 1, m = 0 \rangle - |j = 0, m = 0 \rangle \right) \tag{15.28}$$

▶ **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

$$l \otimes \frac{1}{2} = \left(l + \frac{1}{2}\right) \oplus \left(l - \frac{1}{2}\right) \tag{15.29}$$

Let's construct the states explicitly. First, of course, the top and bottom states of $j = l + 1/2$:

$$\left| j = l + \frac{1}{2}, m = l + \frac{1}{2} \right\rangle = |l, l\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle \tag{15.30}$$

$$\left| j = l + \frac{1}{2}, m = -\left(l + \frac{1}{2}\right) \right\rangle = |l, -l\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \tag{15.31}$$

As usual, we use the lowering operator to obtain the next highest state:

$$\left| j = l + \frac{1}{2}, m = l - \frac{1}{2} \right\rangle = \sqrt{\frac{1}{2l+1}} \, |l, l\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle + \sqrt{\frac{2l}{2l+1}} \, |l, l-1\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle \tag{15.32}$$

Addition of Angular Momentum – States (cont.)

Similarly,

$$
\left| j = l + \frac{1}{2}, m = -\left(l - \frac{1}{2}\right) \right\rangle \tag{15.33}
$$
$$
= \sqrt{\frac{1}{2l+1}}\, |l, -l\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle + \sqrt{\frac{2l}{2l+1}}\, |l, -(l-1)\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle
$$

We determine the top and bottom states of $j = l - 1/2$ by requiring orthogonality, normalization, realness, and the sign-fixing convention:

$$
\left| j = l - \frac{1}{2}, m = l - \frac{1}{2} \right\rangle \tag{15.34}
$$
$$
= \sqrt{\frac{2l}{2l+1}}\, |l, l\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle - \sqrt{\frac{1}{2l+1}}\, |l, l-1\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle
$$

$$
\left| j = l - \frac{1}{2}, m = -\left(l - \frac{1}{2}\right) \right\rangle \tag{15.35}
$$
$$
= \sqrt{\frac{2l}{2l+1}}\, |l, -l\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle - \sqrt{\frac{1}{2l+1}}\, |l, -(l-1)\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle
$$

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:

$$
J_- \left| j = l + \frac{1}{2}, m = n + \frac{1}{2} \right\rangle \tag{15.36}
$$

$$
= \left(L_- \otimes I^{(1/2)} + I^{(l)} \otimes S_- \right) \left(\alpha |l, n+1\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle + \beta |l, n\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle \right)
$$

$$
\hbar \sqrt{\left(l + \frac{1}{2} + n + \frac{1}{2} \right) \left(l + \frac{1}{2} - n - \frac{1}{2} + 1 \right)} \left| j = l + \frac{1}{2}, m = n - \frac{1}{2} \right\rangle \tag{15.37}
$$

$$
= \alpha \hbar \sqrt{(l+n+1)(l-n-1+1)} \, |l, n\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle
$$

$$
+ \beta \hbar \sqrt{(l+n)(l-n+1)} \, |l, n-1\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle + \beta \hbar \, |l, n\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle
$$

which yields

$$
\left| j = l + \frac{1}{2}, m = n - \frac{1}{2} \right\rangle \tag{15.38}
$$
$$
= \left(\alpha \sqrt{\frac{l - n}{l - n + 1}} + \beta \sqrt{\frac{1}{(l + n + 1)(l - n + 1)}} \right) |l, n\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle
$$
$$
+ \beta \sqrt{\frac{l + n}{l + n + 1}} \, |l, n - 1\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle
$$

We have α and β for $n = l - 1$, from these we guess the generic formulae

$$
\alpha_{n=l-1} = \sqrt{\frac{1}{2\,l + 1}} = \sqrt{\frac{l - n}{2\,l + 1}} \qquad \beta_{n=l-1} = \sqrt{\frac{2\,l}{2\,l + 1}} = \sqrt{\frac{l + n + 1}{2\,l + 1}} \tag{15.39}
$$

Inserting these, we find

$$
\left| j = l + \frac{1}{2}, m = n - \frac{1}{2} \right\rangle \tag{15.40}
$$
$$
= \sqrt{\frac{l - n + 1}{2l + 1}} \, |l, n\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle + \sqrt{\frac{l + n}{2l + 1}} \, |l, n - 1\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle
$$

We see that the coefficients obey the expected formulae

$$
\alpha_{n \to n-1} = \sqrt{\frac{l - (n - 1)}{2l + 1}} = \sqrt{\frac{l - n + 1}{2l + 1}} \tag{15.41}
$$

$$
\beta_{n \to n-1} = \sqrt{\frac{l + (n - 1) + 1}{2l + 1}} = \sqrt{\frac{l + n}{2 + 1}} \tag{15.42}
$$

and that the resulting state is correctly normalized. So our guesses for α_n and β_n were correct and the above result for $|j = l + 1/2, m = n - 1/2\rangle$ holds in general for $n = l, \ldots, -l$.

Finally, we can obtain the other state of the same *m* by requiring orthogonality, normalization, realness, and sign-fixing:

$$\left| j = l - \frac{1}{2}, m = n - \frac{1}{2} \right\rangle \tag{15.43}$$
$$= \sqrt{\frac{l+n}{2l+1}} \, |l, n\rangle \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle - \sqrt{\frac{l-n+1}{2l+1}} \, |l, n-1\rangle \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle$$

▶ **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 $V^{(j)}$, we have $\tau^{(0)} = V^{(0)}$ and $\tau^{(1)} = V^{(1)}$. 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 N^n-dimensional space because there are N^n basis elements (*e.g.*, the set of tensors with 1 in exactly of their N^n elements). 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 tensors of rank n can be constructed by taking outer products of lower rank tensors; in particular, the space of tensors of rank n, $\tau^{(n)}$, is a direct product of n copies of $\tau^{(1)}$, the space of vectors. Since $\tau^{(1)} = V^{(1)}$, we thus have

$$\tau^{(n)} = \prod_{k=1}^{n} \tau^{(1)} = \prod_{k=1}^{n} V^{(1)} \tag{15.44}$$

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 $V^{(j)}$ with $j = 0$ to $j = n$; but the trick is to see how many copies of each $V^{(j)}$ we obtain. We can derive it by induction. Suppose

$$\tau^{(n)} = \sum_{j=1}^{n} C_j^{(n)} V^{(j)} \tag{15.45}$$

where $C_j^{(n)}$ indicates the number of copies of $V^{(j)}$ in $\tau^{(n)}$. Then it should be clear that the formula for $\tau^{(n+1)}$ is

$$\tau^{(n+1)} = V^{(1)} \otimes \sum_{j=0}^{n} C_j^{(n)} V^{(j)} \tag{15.46}$$

$$= \sum_{j=1}^{n} C_j^{(n)} \left(V^{(j+1)} \oplus V^{(j)} \oplus V^{(j-1)} \right) \oplus C_0^{(n)} V^{(1)} \tag{15.47}$$

$$= C_n^{(n)} V^{(n+1)} + \left(C_n^{(n)} + C_{n-1}^{(n)} \right) V^{(n)} \tag{15.48}$$

$$\oplus \sum_{j=1}^{n-1} \left(C_{j+1}^{(n)} + C_j^{(n)} + C_{j-1}^{(n)} \right) V^{(j)} \oplus C_1^{(n)} V^{(0)}$$

Summarizing,

$$C_{n+1}^{(n+1)} = C_n^{(n)} \quad C_n^{(n+1)} = C_n^{(n)} + C_{n-1}^{(n)} \quad C_j^{(n+1)} = C_{j+1}^{(n)} + C_j^{(n)} + C_{j-1}^{(n)} \quad C_0^{(n+1)} = C_1^{(n)}$$

$$(15.49)$$

We must specify "initial conditions" for the recursion using $n - 2$, which is

$$\tau^{(2)} = V^{(1)} \otimes V^{(1)} = V^{(2)} \oplus V^{(1)} \oplus V^{(0)} \tag{15.50}$$

So,

$$C_3^{(3)} = 1 \quad C_2^{(3)} = 2 \quad C_1^{(3)} = 3 \quad C_0^{(3)} = 1 \tag{15.51}$$

$$C_4^{(4)} = 1 \quad C_3^{(4)} = 3 \quad C_2^{(4)} = 6 \quad C_1^{(4)} = 6 \quad C_0^{(4)} = 3 \tag{15.52}$$

$$C_5^{(5)} = 1 \quad C_4^{(5)} = 4 \quad C_3^{(5)} = 10 \quad C_2^{(5)} = 15 \quad C_1^{(5)} = 15 \quad C_0^{(5)} = 6 \tag{15.53}$$

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:

$$C_n^{(n)} = 1 \qquad C_{n-1}^{(n)} = n - 1 \tag{15.54}$$

▶ **Spinors?**
For completeness, we note that a construction like the one above for tensors of rank n is unnecessary for spinors. By definition — *i.e.*, because of the way we defined them in Section 14.2 — spin-$n/2$ spinors live in the space $V^{(n/2)}$.

However, one could choose to obtain the spin-$n/2$ spinors by adding the integral angular momentum $j = (n-1)/2$ and $j = 1/2$ to obtain $j = n/2$ and $j = n/2 - 1$:

$$V^{((n-1)/2)} \otimes V^{(1/2)} = V^{(n/2)} \oplus V^{(n/2-1)} \tag{15.55}$$

or adding $j = (n+1)/2$ and $j = 1/2$ to obtain $j = n/2 + 1$ and $j = n/2$

$$V^{((n+1)/2)} \otimes V^{(1/2)} = V^{(n/2+1)} \oplus V^{(n/2)} \tag{15.56}$$

Fortunately, we have already done this in calculating the addition of orbital angular momentum and spin-1/2. So, all that work carries through, and we thus have explicit formulae for constructing any desired $|j = n/2, m\rangle$ state.

We shall see below a method for "direct" definition of spherical tensors of rank n to obtain $V^{(n)}$ without going through the intermediary of $\tau^{(n)}$.

Lecture 42:
Spherical Tensor States and Operators
Selection Rules and the Wigner-Eckart Theorem
Using Tensor Operators to Explain Accidental Degeneracies
Representation Theory

Revision Date: 2008/02/15

Spherical Tensor States

Classical Spherical Tensors and Quantum Mechanical Spherical Tensor States

Recall that we defined classical tensors of rank n in terms of their behavior under a passive rotation via

$$T'_{i_1 \cdots i_n} = \sum_{j_1, j_2, \cdots, j_n} \left(\mathbf{R}_{-\vec{\theta}} \right)_{i_1 j_1} \cdots \left(\mathbf{R}_{-\vec{\theta}} \right)_{i_n j_n} T_{j_1 \cdots j_n} \tag{15.57}$$

where F' is rotated from F by an angle θ CCW about $\widehat{\theta}$; the use of $\mathbf{R}_{-\vec{\theta}}$ (instead of $\mathbf{R}_{\vec{\theta}}$) is consistent with our original definition in Section 14.2 because the transformation of the tensor is *passive* – it is the same tensor, but in a different coordinate system. We know that the rotation matrices can be obtained from the $j = 1$ angular momentum generators (see, *e.g.*, Section 13.4)

$$\mathbf{R}_{-\vec{\theta}} = \exp\left(-\vec{\theta} \cdot \vec{\mathbf{M}} \right) = \exp\left(-\frac{i}{\hbar} \left(-\vec{\theta} \right) \cdot i \hbar \vec{\mathbf{M}} \right) \tag{15.58}$$

This definition takes a fundamentally *Cartesian* point of view because it relates coordinate representations that are the projections of the tensor onto Cartesian unit vectors.

Spherical Tensor States (cont.)

On the other hand, for the quantum-mechanical spin-j spinor states (j half-integral), we defined them by fiat to be objects $|\psi^{(j)}\rangle$ consisting of $2j+1$ elements $\{|\psi_q^{(j)}\rangle\}$, $q = 1, \ldots, 2j+1$ (corresponding to the coefficient of $|j, m_q\rangle$ with $m_q = j - q + 1$ in an expansion in the $|j, m\rangle$ basis for this j) that transform under an *active* rotation transformation by an angle $\vec{\theta}$ by

$$|\psi^{(j)\,\prime}\rangle = \exp\left(-\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{J}^{(j)}\right)|\psi^{(j)}\rangle \tag{15.59}$$

or, component-wise,

$$|\psi_p^{(j)\,\prime}\rangle = \sum_{q=1}^{2j+1}\left[\exp\left(-\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{J}^{(j)}\right)\right]_{pq}|\psi_q^{(j)\,\prime}\rangle \tag{15.60}$$

Remember that, for QM tensor states, it was necessary to always talk in terms of *active* transformations — *i.e.*, physically rotate the state relative to the underlying physical space — in order to avoid being tied to the position basis. The $2j+1$ index values q correspond to the coefficients of the state $|j, m_q\rangle$ with $m_q = j - q + 1$ assuming we order the states $|j, j\rangle, |j, j-1\rangle, \ldots, |j, -j\rangle$.

We may infer a definition of a *classical* spin-j spinor from the above: a classical spin-j spinor $\vec{v}^{(j)}$ is an object that has coordinate representations $\underline{v}^{(j)}$ consisting of $2j + 1$ components $\{v_q^{(j)}\}$, $q = 1, \ldots, 2j + 1$, that transform under a passive coordinate system transformation (F' obtained from F by rotation by $\vec{\theta}$, but we want to calculate the components of $\vec{v}$ in F' from those in F; we are not seeking to *rotate* $\vec{v}^{(j)}$) by

$$\underline{v}^{(j)\,\prime} = \exp\left(\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{J}^{(j)}\right)\underline{v}^{(j)} \tag{15.61}$$

or, component-wise,

$$v_p^{(j)\,\prime} = \sum_{q=1}^{2j+1}\left[\exp\left(\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{J}^{(j)}\right)\right]_{pq} v_q^{(j)} \tag{15.62}$$

Here, we have used a *passive* rotation because that is the natural way to define tensors classically; one essentially *always* works in a "position basis" in classical mechanics, so being tied to that position basis is not problematic.

We thus recognize that we could have simply defined, by fiat in the same way as we do spinors, classical objects $\mathcal{T}^{(j)}$ having coordinate representations $\underline{\mathcal{T}}_q^{(j)}$ consisting of $2j + 1$ components $\{T_q^{(j)}\}$ (again, $q = 1, \ldots, 2j + 1$) that are related under a *passive* transformation (again, F' obtained from F by rotation by $\vec{\theta}$, but we want to calculate the components of $\mathcal{T}^{(j)}$ in F' from those in F; we are not seeking to *rotate* $\mathcal{T}^{(j)}$) by

$$\underline{\mathcal{T}}^{(j)\,\prime} = \exp\left(\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{J}^{(j)}\right)\underline{\mathcal{T}}^{(j)} \tag{15.63}$$

or, component-wise,

$$T_p^{(j)\,\prime} = \sum_{q=1}^{2j+1}\left[\exp\left(\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{J}^{(j)}\right)\right]_{pq} T_q^{(j)} \tag{15.64}$$

These *classical* objects are termed spherical tensors of rank j where j is integral. With this definition, we see that spinors are also classical spherical tensors, but with half-integral rank j.

Finally, let us define quantum mechanical spherical tensor states of rank *j* by simple extension of the definition of half-integral *j* spinor states: a spherical tensor state of rank *j* is an object $|\psi^{(j)}\rangle$ that has $2j+1$ elements $\{|\psi_q^{(j)}\rangle\}$, $q = 1, \ldots, 2j+1$ (again, corresponding to the coefficient of $|j, m_q\rangle$ with $m_q = j - q + 1$ in an expansion in the $|j, m\rangle$ basis for this *j*) that transform under an *active* rotation transformation by

$$|\psi^{(j)}{}'\rangle = \exp\left(-\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{J}^{(j)}\right)|\psi^{(j)}\rangle \tag{15.65}$$

or, component-wise,

$$|\psi_p^{(j)}{}'\rangle = \sum_{q=1}^{2j+1}\left[\exp\left(-\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{J}^{(j)}\right)\right]_{pq}|\psi_q^{(j)}\rangle \tag{15.66}$$

Again, remember that for QM tensor states, we must always talk in terms of active transformations to avoid being tied to the position basis. Again, the $2j+1$ index values *q* correspond to the coefficients of the state $|j, j - q + 1\rangle$ assuming we order the states $|j, j\rangle, |j, j-1\rangle, \ldots, |j, -j\rangle$. The spherical tensors states of rank *j* are simply the elements of $V^{(j)}$, and hence we will from now on refer to those spaces not as "angular momentum *j*" spaces but as spaces of "spherical tensors states of rank *j*."

Connecting Spherical and Cartesian Tensors

What is the algebraic connection of classical spherical tensors of rank j and classical Cartesian tensors of rank j? That is, how does one convert between them? We must recognize based on the breakdown of $\tau^{(j)}$ into a direct sum of $V^{(j)}$ that the spherical tensors of rank j are only a subspace of the Cartesian tensors of rank j, as is evidenced by simply counting the dimensionality, N^j vs. $2j + 1$. But the breakdown also leads us to recognize that a set of generalized Clebsch-Gordan coefficients connects the two representations.

Spherical Tensor States (cont.)

In detail: we have a Cartesian basis for the Cartesian tensors of rank n

$$\varepsilon_{i_1 \cdots i_n} = \prod_{k=1}^{n} \vec{e}_{i_k} \tag{15.67}$$

where the $\{\vec{e}_{i_k}\}$ are unit vectors in the i_kth direction and hence belong to $\tau^{(1)}$, the space of rank 1 Cartesian tensors. $\prod$ is a multiple direct product. Recognizing that $\tau^{(1)} = V^{(1)}$, we may instead choose as a basis for $\tau^{(1)}$ the eigenstates of $i\hbar \mathbf{M}_z$, which we found earlier to be $\vec{v}_{\pm 1} = [1, \pm i, 0]^T$ and $\vec{v}_0 = [0, 0, 1]^T$. That is, define our basis to be

$$\prod_{k=1}^{n} \vec{v}_{i_k} = \prod_{k=1}^{n} |j_k = 1, i_k \rangle \tag{15.68}$$

where $i_k = \pm 1, 0$. Our basis now consists of direct products of $j = 1$ states of particular values of m! This is just an uncoupled basis from the point of view of angular momentum addition.

There is a corresponding coupled basis, which consists of the natural basis elements of the direct sum representation of this direct product space, and there will be generalized Clebsch-Gordan coefficients connecting the two:

$$|j_i, m\rangle = \sum_{m_k = 0, \pm 1} \left(\prod_{k=1}^{n} |j_k = 1, m_k\rangle \right) \left(\prod_{k=1}^{n} \langle j_k = 1, m_k | \right) |j_i, m\rangle \qquad (15.69)$$

where $\left(\prod_{k=1}^{n} \langle j_k = 1, m_k | \right) |j_i, m\rangle$ are the Clebsch-Gordan coefficients. There is a i index on j_i because there may be multiple orthogonal copies of the same $V(j)$ subspace in the direct sum; i indexes them. Of course, it will be tremendously tedious to calculate all these coefficents for $n > 2$, but it is feasible. There are actually 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.

We have this nice conversion from Cartesian to spherical tensors for integral rank j, in which the space of spherical tensors of rank j, $V^{(j)}$, is a subspace of the space of Cartesian tensors of rank j, $\tau^{(j)}$ (or of $\tau^{(n)}$ for $n > j$). Is there a similar set of Cartesian tensors for half-integral j? Sort of.

It is hard to think of a way to represent spherical tensors of half-integral rank in a completely Cartesian fashion because one would need a half-integral number of rotation matrices! But one can of course obtain a spherical tensor of half-integral rank $j > 1/2$ as a subspace of the direct product of Cartesian tensors of rank $j - 1/2$ and spherical tensors of rank $1/2$ (or of $j + 1/2$ and $1/2$). But one fundamentally needs to include one space of spherical tensors of half-integral rank from the start to get spherical tensors of half-integral rank as a subspace of the result, so the representation is not completely Cartesian in character.

Which one to use, Cartesian or spherical?

In the end though, since we can break down any Cartesian tensor state of rank n into spherical tensor states of rank n and lower, we are free to define spherical tensor states to start off with instead of defining Cartesian tensor states. Which one one wants depends on the situation. If the Hamiltonian for a particle described by tensor states is most naturally expressed in spherical coordinates, then spherical tensors are a good starting point. If the Hamiltonian depends most strongly on Cartesian coordinates, then Cartesian tensors make most sense.

For example, a spherically symmetric potential naturally leads one to spherical tensors states since they will immediately be eigenstates of H and the degeneracies of H will naturally be explained in terms of its dependence only on J^2 as opposed to depending on J_z. A non-spherically symmetric potential, such as an anisotropic SHO, suggests Cartesian tensor states.

Spherical Tensor Operators

Defining Spherical Tensor Operators

Recall that we defined a (Cartesian) tensor operator of rank n, $\mathcal{O}$, to be a set of operators that, under the unitary transformation corresponding to a coordinate system rotation, are linearly combined in the same way as a classical coordinate transformation for a tensor of rank n would mix them:

$$O'_{i_1 \cdots i_n} = T(\vec{\theta}) O_{i_1 \cdots i_n} T^{\dagger}(\vec{\theta}) = \sum_{j_1, \cdots, j_n = 1}^{N} R^T_{i_1 j_1} \cdots R^T_{i_n j_n} O_{j_1 \cdots j_n} \tag{15.70}$$

where $T(\vec{\theta}) = \exp\left(-\frac{i}{\hbar} \vec{\theta} \cdot \vec{J}\right)$, $R_{ij} = \left(\mathbf{R}_{\vec{\theta}}\right)_{ij}$, and where here we are writing the representations of the tensor operator $\mathcal{O}$ of rank n as $O_{i_1 \cdots i_n}$ in frame F and $O'_{i_1 \cdots i_n}$ in frame F', where F' is rotated by $\vec{\theta}$ from F. $\vec{J}$ operates in whatever space T does, whereas $\mathbf{R}$ is always just a $N \times N$ matrix of numbers. The fact that $\mathbf{R}$ must be transposed was derived in Section 14.2 and recalls the definition of a *classical* Cartesian tensor. (Actually, we never explicitly wrote the above equation for arbitrary rank n, but we did write it for vector operators.)

It should be obvious that one can define spherical tensor operators in the same way as we did classical spherical tensors. That is, a spherical tensor operator of rank j, $\mathcal{O}^{(j)}$, is a set of operators whose representations in different coordinate systems (coordinate representations, just as for Cartesian tensors) $\underline{\mathcal{O}}^{(j)}$, consisting of $2j + 1$ components $\{O_q^{(j)}\}$, $q = 1, \ldots, 2j + 1$, are related by

$$O_q^{(j)\,\prime} = T(\vec{\theta})\,O_q^{(j)}\,T^{\dagger}(\vec{\theta}) = \sum_{p=1}^{2j+1} \left[\exp\left(-\frac{i}{\hbar}\,\vec{\theta} \cdot \vec{J}^{(j)} \right) \right]_{pq} O_p^{(j)} \qquad (15.71)$$

where it is understood that $\underline{\mathcal{O}}^{(j)}$ is to be treated as a column vector with $2j + 1$ components indexed by q from 1 to $2j + 1$: on the right side, $\vec{J}^{(j)}$ is a $(2j + 1) \times (2j + 1)$ matrix, as is its exponent, that acts on the column vector (note the transposition, though!); and on the left side, the $T(\vec{\theta})$ and $T^{\dagger}(\vec{\theta})$ operators act on each component of the column vectors separately. The $\vec{J}^{(j)}$ matrix is the *same* matrix we use on spherical tensor states of rank j. The fact that the rotation operator on the right side acts to the left on $O_p^{(j)}$ recalls the passive nature of the transformation defining classical spherical tensors; action of this rotation operator to the left is like action of the adjoint operator (corresponding to an opposite sign for the rotation angle) to the right.

We can write the above in more compact form as

$$\underline{\mathcal{O}}^{(j)\,\prime} = T(\vec{\theta})\,\underline{\mathcal{O}}^{(j)}\,T^\dagger(\vec{\theta}) = \underline{\mathcal{O}}^{(j)} \exp\left(-\frac{i}{\hbar}\,\vec{\theta}\cdot\vec{J}^{(j)} \right) \tag{15.72}$$

We could not write the definition for Cartesian tensors in this compact form because there was no compact way to write the action of the multiple rotation matrices.

We note that this definition also allows for spherical tensor operators of half-integral rank just as the analogous definition allowed classical spherical tensors of half-integral rank. One must be careful, however, not to confuse the rank of an operator and the rank of the states it operates on. For example, $\vec{J}^{(1/2)}$ operates on spherical tensor states of rank $1/2$, but $\vec{J}^{(1/2)}$ is itself a (Cartesian or spherical) tensor operator of rank 1!

Spherical Tensor Operators (cont.)

Examples of Spherical Tensor Operators

▶ **Spherical vector Operators**
Because $\tau^{(1)} = V^{(1)}$ — the space of Cartesian tensor operators of rank 1 is identical to that of spherical tensor operators of rank 1 — any vector operator can be written as a spherical tensor operator. We know the eigenvectors of the $i\,\hbar\,\mathbf{M}_z$ operator (see Section 14.2), which defines spherical tensor states of rank 1, so we can just use those to calculate the components of a spherical tensor operator from the components of the Cartesian tensor operator. Recall that, in the basis of the Cartesian unit vectors, the spherical tensor states are

$$\vec{v}_{+1} = -\frac{1}{\sqrt{2}} \left[\begin{array}{c} 1 \\ i \\ 0 \end{array} \right] \qquad \vec{v}_0 = \left[\begin{array}{c} 0 \\ 0 \\ 1 \end{array} \right] \qquad \vec{v}_{-1} = \frac{1}{\sqrt{2}} \left[\begin{array}{c} 1 \\ -i \\ 0 \end{array} \right] \qquad (15.73)$$

(The $-$ sign in front of the expression for $\vec{v}_{+1}$ is a change from the sign conventions used when deriving these eigenvectors in Section 14.2. The sign convention is set to match that of the Y_l^m.)

Therefore, to project out from a vector operator $\vec{O}$ the elements of the spherical tensor operator $\mathcal{O}^{(1)}$, we take the appropriate inner products:

$$O_1^{(1)} = \langle \vec{O} \,|\, \vec{v}_{+1} \rangle = \begin{bmatrix} O_1 & O_2 & O_3 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ -i \\ 0 \end{bmatrix} = -\frac{O_1 + i\,O_2}{\sqrt{2}} \quad (15.74)$$

$$O_2^{(1)} = \langle \vec{O} \,|\, \vec{v}_0 \rangle = \begin{bmatrix} O_1 & O_2 & O_3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = O_3 \quad (15.75)$$

$$O_3^{(1)} = \langle \vec{O} \,|\, \vec{v}_{-1} \rangle = \begin{bmatrix} O_1 & O_2 & O_3 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \\ 0 \end{bmatrix} = \frac{O_1 - i\,O_2}{\sqrt{2}} \quad (15.76)$$

Why do we take inner products $\langle \vec{O} \,|\, \vec{v}_m \rangle$ instead of $\langle \vec{v}_m \,|\, \vec{O} \rangle$? Recall that, in the transformation rule for a spherical tensor operator $\mathcal{O}^{(j)}$, the transformation rule is $\underline{\mathcal{O}}^{(j)\,\prime} = \underline{\mathcal{O}}^{(j)} \exp\left(-\frac{i}{\hbar}\,\vec{\theta} \cdot \vec{J}^{(j)} \right)$; that is, $\underline{\mathcal{O}}^{(j)}$ is treated as a row vector. So, we needed to project $\vec{O}$ onto $\vec{v}_m^{\dagger}$, not onto $\vec{v}_m$, hence the reversal of the position of $\vec{O}$ and $\vec{v}_m$ in the inner product. This point is very important because the results would be complex conjugated, exchanging $O_1^{(1)}$ and $O_3^{(1)}$, if we had not done this correctly.

▶ **Spherical tensor Operators**
An obvious way to obtain a tensor operator is to take the outer product of two vectors. One such Cartesian tensor operator is the inertia tensor of classical mechanics. The operator definition of this Cartesian tensor is

$$M_{ij} = \delta_{ij} \sum_k R_k R_k - R_i R_j \qquad (15.77)$$

How does *M* break down into spherical tensors? Recall that, in connecting Cartesian and spherical tensors, our approach was to write a basis for Cartesian tensors of rank *n* that consisted of direct products of spherical tensors of rank 1 rathan than of Cartesian unit vectors of rank 1. Once we have that decomposition, we can write these direct products in terms of spherical tensors of rank 2, 1, and 0 using our angular momentum addition rules (Clebsch-Gordan coefficients). Proceeding step-by-step:

1. **Writing down a basis for Cartesian tensors of rank 2 in terms of direct products of spherical tensors of rank 1.**
   This is easy, we just need to take all possible direct products of the three spherical tensors of rank 1 that map directly to spherical tensor unit vectors. That is,

$$|1, \pm 1 \rangle \otimes |1, \pm 1 \rangle \leftrightarrow \vec{v}_{\pm 1} \vec{v}_{\pm 1}^T = \frac{1}{2} \begin{bmatrix} 1 & \pm i & 0 \\ \pm i & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{15.78}$$

$$|1, \pm 1 \rangle \otimes |1, \mp 1 \rangle \leftrightarrow \vec{v}_{\pm 1} \vec{v}_{\mp 1}^T = -\frac{1}{2} \begin{bmatrix} 1 & \mp i & 0 \\ \pm i & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{15.79}$$

$$|1, \pm 1 \rangle \otimes |1, 0 \rangle \leftrightarrow \vec{v}_{\pm 1} \vec{v}_0^T = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & \mp 1 \\ 0 & 0 & -i \\ 0 & 0 & 0 \end{bmatrix} \tag{15.80}$$

$$|1, 0 \rangle \otimes |1, \pm 1 \rangle \leftrightarrow \vec{v}_0 \vec{v}_{\pm 1}^T = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \mp 1 & -i & 0 \end{bmatrix} \tag{15.81}$$

$$|1, 0 \rangle \otimes |1, 0 \rangle \leftrightarrow \vec{v}_0 \vec{v}_0^T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{15.82}$$

Why a ^{T} instead of a † on the second element of each outer product? How would you decide which one to put the * on?

To explain the above basis, consider an example: a Cartesian tensor operator with all elements except O_{33} vanishing in some coordinate system would be equivalent to O_{33} ($|1, 0\rangle \otimes |1, 0\rangle$).

2. **Write the spherical tensors of rank 2, 1, and 0 in terms of the direct products.**
We just need Clebsch-Gordan coefficients. We pull them from tables (e.g., PDG), giving:

$$|2, \pm 2\rangle = |1, \pm 1\rangle \otimes |1, \pm 1\rangle \tag{15.83}$$

$$|2, \pm 1\rangle = \sqrt{\frac{1}{2}} |1, \pm 1\rangle \otimes |1, 0\rangle + \sqrt{\frac{1}{2}} |1, 0\rangle \otimes |1, \pm 1\rangle \tag{15.84}$$

$$|2, 0\rangle = \sqrt{\frac{1}{6}} |1, 1\rangle \otimes |1, -1\rangle + \sqrt{\frac{2}{3}} |1, 0\rangle \otimes |1, 0\rangle + \sqrt{\frac{1}{6}} |1, -1\rangle \otimes |1, 1\rangle \tag{15.85}$$

$$|1, \pm 1\rangle = \pm \left[\sqrt{\frac{1}{2}} |1, \pm 1\rangle \otimes |1, 0\rangle - \sqrt{\frac{1}{2}} |1, 0\rangle \otimes |1, \pm 1\rangle \right] \tag{15.86}$$

$$|1, 0\rangle = \sqrt{\frac{1}{2}} |1, 1\rangle \otimes |1, -1\rangle - \sqrt{\frac{1}{2}} |1, -1\rangle \otimes |1, 1\rangle \tag{15.87}$$

$$|0, 0\rangle = \sqrt{\frac{1}{3}} |1, 1\rangle \otimes |1, -1\rangle - \sqrt{\frac{1}{3}} |1, 0\rangle \otimes |1, 0\rangle + \sqrt{\frac{1}{3}} |1, -1\rangle \otimes |1, 1\rangle \tag{15.88}$$

Finally, combining the two, and taking the necessary adjoint of O,

$$O_{3\mp2}^{(2)} = \langle O|2,\pm2\rangle = \langle O|(|1,\pm1\rangle \otimes |1,\pm1\rangle) \tag{15.89}$$

$$= \sum_{i_1,i_2=1}^{3} O_{i_1 i_2}\left(\vec{v}_{\pm1}\vec{v}_{\pm1}^T\right)_{i_1 i_2} = \frac{1}{2}\left[O_{11} \pm i\,O_{12} \pm i\,O_{21} - O_{22}\right] \tag{15.90}$$

$$O_{3\mp1}^{(2)} = \langle O|2,\pm1\rangle = \frac{1}{2}\left[\mp O_{13} \mp O_{31} - i\,O_{23} - i\,O_{32}\right] \tag{15.91}$$

$$O_3^{(2)} = \langle O|2,0\rangle = \sqrt{\frac{1}{6}}\left[-O_{11} - O_{22} + O_{33}\right] \tag{15.92}$$

$$O_{2\mp1}^{(1)} = \langle O|1,\pm1\rangle = \frac{1}{2}\left[-O_{13} + O_{31} \mp i\,O_{23} \pm i\,O_{32}\right] \tag{15.93}$$

$$O_2^{(1)} = \langle O|1,0\rangle = \sqrt{\frac{1}{2}}\left[i\,O_{12} - i\,O_{21}\right] \tag{15.94}$$

$$O_1^{(0)} = \langle O|0,0\rangle = \sqrt{\frac{1}{3}}\left[-O_{11} - O_{22} - O_{33}\right] \tag{15.95}$$

So, finally, we may calculate the moment of inertia tensor in spherical tensor form. We have

$$M_{3\pm2}^{(2)} = \frac{1}{2}\left[\left(\vec{R}\cdot\vec{R}\right)(1-1)\right] - \frac{1}{2}\left[(R_1 R_1 \pm i R_1 R_2 \pm i R_2 R_1 - R_2 R_2)\right] \quad (15.96)$$

$$= -\frac{1}{2}\left(R_1^2 - R_2^2 \pm 2i R_1 R_2\right) \quad (15.97)$$

$$M_{3\pm1}^{(2)} = -\frac{1}{2}\left[\mp R_1 R_3 \mp R_3 R_1 - i R_2 R_3 - i R_3 R_2\right] = \pm R_1 R_3 + i R_2 R_3 \quad (15.98)$$

$$M_3^{(2)} = \sqrt{\frac{1}{6}}\left\{\left[\vec{R}\cdot\vec{R}\right](-1-1+1)\right] - \left[-R_1^2 - R_2^2 + R_3^2\right]\right\} = -\sqrt{\frac{2}{3}}\,R_3^2 \quad (15.99)$$

$$M_{2\pm1}^{(1)} = -\frac{1}{2}\left[-R_1 R_3 + R_3 R_1 \mp i R_2 R_3 \pm i R_3 R_2\right] = 0 \quad (15.100)$$

$$M_2^{(1)} = -\sqrt{\frac{1}{2}}\left[i R_1 R_2 - i R_2 R_1\right] = 0 \quad (15.101)$$

$$M_1^{(0} = \sqrt{\frac{1}{3}}\left\{\left[-\vec{R}\cdot\vec{R}(1+1+1)\right] - \left[-R_1 R_1 - R_2 R_2 - R_3 R_3\right]\right\} = -\frac{2}{\sqrt{3}}\,\vec{R}\cdot\vec{R} \quad (15.102)$$

The result makes physical sense: the inertia tensor is symmetric, so it only has 6 independent components. One of these is the trace, which is known to be a

scalar under rotations (the trace of a second rank tensor is invariant under rotations, so the trace ($\vec{R} \cdot \vec{R}$) comes out as the rank 0 spherical tensor. The remaining 5 components map to the 5 rank 2 spherical tensor components.

Properties of Spherical Tensor Operators

You are no doubt wondering how spherical tensor operators are connected to Cartesian tensor operators; we will answer that question, but we first need to prove a few interesting properties of these operators. These will make use of the defining relation between the action of the unitary transformation corresponding to a coordinate system rotation and the action of the "classical" coordinate system rotation on the operator.

▶ **Irreducibility**
The first thing to show is that a spherical tensor operator of rank j cannot be written as a direct sum of spherical tensor operators of lower rank; it is therefore considered **irreducible**. This follows immediately from the defining relation: a spherical tensor operator of rank j transforms under a coordinate system rotation in the *exact same way* as a state in the space $V^{(j)}$. We have already shown that the space $V^{(j)}$ is irreducible because the raising and lowering operators connect every $|j, m\rangle$ basis element of $V^{(j)}$ to every other one. That means that any state in $V^{(j)}$ cannot be written in terms of basis elements of states in $V^{(k)}$ for $k < j$. This, in turn, implies that there is no way to simplify the transformation rules for a state in $V^{(j)}$ to use operators that act in $V^{(k)}$ for $k < j$. Hence, the same is true for our spherical tensor operator of rank j.

▶ **Action on $|j, m\rangle$ states**

Let's consider the action of a spherical tensor operator of rank k $\mathcal{O}^{(k)}$ on an angular momentum eigenstate $|j, m\rangle$. The way to figure this out is to determine what properties $\mathcal{O}_q^{(k)}|j, m\rangle$ has under rotations. We simply apply an active rotation transformation to this state:

$$T(\vec{\theta})\left[O_q^{(k)}|j, m\rangle\right] = \left(T(\vec{\theta})O_q^{(k)}T^{\dagger}(\vec{\theta})\right)\left(T(\vec{\theta})|j, m\rangle\right) \tag{15.103}$$

Evaluating the second factor is actually not straightforward, let's do it separately (with: $m_a \equiv j - a + 1$ and $m_b \equiv j - b + 1$):

$$T(\vec{\theta})|j, m\rangle = \sum_{a,b=1}^{2j+1} |j, m_a\rangle\langle j, m_a | T(\vec{\theta})|j, m_b\rangle\langle j, m_b |j, m\rangle \tag{15.104}$$

$$= \sum_{a,b=1}^{2j+1} |j, m_a\rangle \left[\exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot\vec{J}^{(j)}\right)\right]_{ab} \delta_{m_b,m} \tag{15.105}$$

$$= \sum_{a=1}^{2j+1} \left[\exp\left(-\frac{i}{\hbar}\vec{\theta}\cdot\vec{J}^{(j)}\right)\right]_{a,j-m+1} |j, m_a\rangle \tag{15.106}$$

Now, let's evaluate the whole thing, using the definition of a spherical tensor operator for the first term. We have:

$$T(\vec{\theta}) \left[O_q^{(k)} |j, m\rangle \right] \tag{15.107}$$
$$= \left[\sum_{p=1}^{2k+1} \left[\exp\left(-\frac{i}{\hbar} \vec{\theta} \cdot \vec{J}^{(k)} \right) \right]_{pq} O_p^{(k)} \right] \left[\sum_{a=1}^{2j+1} \left[\exp\left(-\frac{i}{\hbar} \vec{\theta} \cdot \vec{J}^{(j)} \right) \right]_{a,j-m+1} |j, m_a\rangle \right]$$

We see that the action of the transformation on $\mathcal{O}_q^{(k)}$ is the same as its action on $|j, m\rangle$, once one sorts out the indices (q is like $j - m + 1$: $q = 1, \ldots, 2j+1$ corresponds to $m = j, \ldots, -j$). Moreover, we see that $\mathcal{O}_q^{(k)} |j, m\rangle$ transforms like the product state $|k, m_q\rangle \otimes |j, m\rangle$ (with $m_q \equiv k - q + 1$, $m_p \equiv k - p + 1$):

$$T(\vec{\theta}) \left[|k, m_q\rangle \otimes |j, m\rangle \right] \tag{15.108}$$
$$= \left[\sum_{p=1}^{2k+1} \left[\exp\left(-\frac{i}{\hbar} \vec{\theta} \cdot \vec{J}^{(k)} \right) \right]_{pq} |k, m_p\rangle \right] \otimes \left[\sum_{a=1}^{2j+1} \left[\exp\left(-\frac{i}{\hbar} \vec{\theta} \cdot \vec{J}^{(j)} \right) \right]_{a,j-m+1} |j, m_a\rangle \right]$$

Since the two states transform in the same way, they must be the same (up to normalization):

$$O_q^{(k)}|j, m\rangle = |k, m_q\rangle \otimes |j, m\rangle \tag{15.109}$$

That is, acting on the spherical tensor state $|j, m\rangle$ with the qth component of a spherical tensor operator of rank k yields the product state $|k, m_q\rangle \otimes |j, m\rangle$. Spherical tensor operators add angular momentum.

Spherical Tensor Operators (cont.)

▶ **Commutation Relations**
We can prove the commutation relations

$$\left[J_\pm, O_q^{(j)} \right] = \pm \hbar \sqrt{(j \mp m_q)(j \pm m_q + 1)}\, O_{q \mp 1}^{(j)} \tag{15.110}$$

$$\left[J_z, O_q^{(j)} \right] = \hbar\, m_q\, O_q^{(j)} \tag{15.111}$$

significance will become clear below. Recall $q = 1, \ldots, 2j + 1$ and
$m_q = j - q + 1$.

We do the proofs by writing down the definition of the rotation properties of a spherical tensor for an infinitesimal rotation by $\delta\theta$ about an arbitrary axis $\widehat{n}$:

$$T(\delta\theta\,\widehat{n})\,\mathcal{O}^{(j)}\,T^{\dagger}(\delta\theta\,\widehat{n}) = \mathcal{O}^{(j)}\exp\left(-\frac{i}{\hbar}\,\delta\theta\,\widehat{n}\cdot\vec{J}^{(j)}\right) \quad (15.112)$$

$$\left(1 - \frac{i}{\hbar}\,\delta\theta\,\widehat{n}\cdot\vec{J}\right)\mathcal{O}^{(j)}\left(1 + \frac{i}{\hbar}\,\delta\theta\,\widehat{n}\cdot\vec{J}\right) = \mathcal{O}^{(j)}\left(1 - \frac{i}{\hbar}\,\delta\theta\,\widehat{n}\cdot\vec{J}^{(j)}\right) \quad (15.113)$$

$$\left[\widehat{n}\cdot\vec{J}, \mathcal{O}^{(j)}\right] = \mathcal{O}^{(j)}\widehat{n}\cdot\vec{J}^{(j)} \quad (15.114)$$

One must be very careful to recognize that the $\widehat{n}\cdot\vec{J}$ operator on the left side is an operator in the same space that $\mathcal{O}^{(j)}$ operates on, while the $\widehat{n}\cdot\vec{J}^{(j)}$ object on the right side is just a matrix that operates on column vectors with $2j+1$ elements. It is very easy to confuse them and think they are the same object. They are of course related, but, for the purposes of the above equation, they are distinct objects. Component-wise, the above becomes

$$\left[\widehat{n}\cdot\vec{J}, O_q^{(j)}\right] = \sum_{p=1}^{2j+1}\left[\widehat{n}\cdot\vec{J}^{(j)}\right]_{pq}O_p^{(j)} \quad (15.115)$$

Spherical Tensor Operators (cont.)

Now, let us consider special cases in order to obtain the desired commutation relations. First, take $\widehat{n} = \widehat{z}$:

$$\left[J_z, O_q^{(j)} \right] = \sum_{p=1}^{2j+1} \left[J_z^{(j)} \right]_{qp} O_p^{(j)} = \sum_{p=1}^{2j+1} m_q \, \hbar \, \delta_{qp} \, O_p^{(j)} \tag{15.116}$$

$$= m_q \, \hbar \, O_q^{(j)} \tag{15.117}$$

where we used the known diagonal structure of the $J_z^{(j)}$ matrix. Similarly,

$$\left[J_\pm, O_q^{(j)} \right]$$

$$= \left[J_x \pm i \, J_y, O_q^{(j)} \right] = \sum_{p=1}^{2j+1} \left[J_x^{(j)} \pm i \, J_y^{(j)} \right]_{pq} O_p^{(j)} = \sum_{p=1}^{2j+1} \left[J_\pm^{(j)} \right]_{pq} O_p^{(j)} \tag{15.118}$$

$$= \sum_{p=1}^{2j+1} \pm \hbar \sqrt{(j \mp m_q)(j \pm m_q + 1)} \, \delta_{p \pm 1, q} \, O_p^{(j)} \tag{15.119}$$

$$= \pm \hbar \sqrt{(j \mp m_q)(j \pm m_q + 1)} \, O_{q \mp 1}^{(j)} \tag{15.120}$$

▶ **Selection Rules**
Because of the way in which spherical tensor operators add angular momentum, we are able to make strong statements about the matrix elements of a spherical tensor operator:

$$\langle j', m' | O_q^{(k)} | j, m \rangle = 0 \qquad \text{unless} \quad |k - j| \le j' k + j, \quad m' = m + m_q \quad (15.121)$$

That is, unless the state on the left side has a nonzero projection onto the state obtained by adding angular momenta $|j, m\rangle$ and $|k, m_q\rangle$, then the matrix element vanishes. This follows immediately from the proof that $O_q^{(k)} |j, m\rangle \propto |k, m_q\rangle \otimes |j, m\rangle$. This kind of requirement is known as a *selection rule* and will be very useful when we consider perturbation theory.

Spherical Tensor Operators (cont.)

▶ **Wigner-Eckart Theorem**
The Wigner-Eckart Theorem provides a means to simplify the matrix elements of a spherical tensor operator that are not shown to vanish via selection rules. Given a tensor operator $O_q^{(k)}$, it states

$$\langle \alpha', j', m' | O_q^{(k)} | \alpha, j, m \rangle = \langle \alpha', j' || \mathcal{O}^{(k)} || \alpha, j \rangle \langle j', m' | (|k, m_q\rangle \otimes |j, m\rangle) \tag{15.122}$$

where $m_q = k - q + 1$ and α and α' indicate any other quantum numbers in the problem. The second term carries all the angular information and enforces selection rules. The first term carries information that does not depend on m, m', or m_q. Clearly, this can prove to be a great simplification in evaluating matrix elements without having to do lots of solid angle integrals.

The proof is shockingly simple. Our commutation relation for $[J_\pm, O_q^{(k)}]$ implies

$$\langle \alpha', j', m' | \left[J_\pm, O_q^{(k)} \right] | \alpha, j, m \rangle$$
$$= \hbar \sqrt{(k \mp m_q)(k \pm m_q + 1)} \, \langle \alpha', j', m' | O_{q\mp1}^{(k)} | \alpha, j, m \rangle \tag{15.123}$$

Let's evaluate the left-hand side using the standard action of $J_\pm$ on the states:

$$
\begin{aligned}
\langle \alpha', j', m' | &\left[J_\pm, O_q^{(k)} \right] | \alpha, j, m \rangle \\
&= \hbar \sqrt{(j' \pm m')(j' \mp m' + 1)} \, \langle \alpha', j', m' \mp 1 | O_q^{(k)} | \alpha, j, m \rangle \\
&\quad - \sqrt{(j \mp m)(j \pm m + 1)} \, \langle \alpha', j', m' | O_q^{(k)} | \alpha, j, m \pm 1 \rangle
\end{aligned}
\tag{15.124}
$$

Combining the two, we have

$$
\begin{aligned}
\sqrt{(j' \pm m')(j' \mp m' + 1)} \, &\langle \alpha', j', m' \mp 1 | O_q^{(k)} | \alpha, j, m \rangle \\
&= \sqrt{(k \mp m_q)(k \pm m_q + 1)} \, \langle \alpha', j', m' | O_{q \mp 1}^{(k)} | \alpha, j, m \rangle \\
&\quad + \sqrt{(j \mp m)(j \pm m + 1)} \, \langle \alpha', j', m' | O_q^{(k)} | \alpha, j, m \pm 1 \rangle
\end{aligned}
\tag{15.125}
$$

Spherical Tensor Operators (cont.)

Now, we may derive a similar relation for Clebsch-Gordan coefficients:

$$\hbar\sqrt{(j' \pm m')(j' \mp m' + 1)} \langle j', m' \mp 1 \,|\, (|k, m_q\rangle \otimes |j, m\rangle)$$
$$= \langle j', m' | J_\pm (|k, m_q\rangle \otimes |j, m\rangle) \qquad (15.126)$$
$$= \langle j', m' | \left(J_\pm^{(1)} \otimes I^{(2)} + I^{(1)} \otimes J_\pm^{(2)} \right) (|k, m_q\rangle \otimes |j, m\rangle) \qquad (15.127)$$
$$= \hbar\sqrt{(k \mp m_q)(k \pm m_q + 1)} \langle j', m' | (|k, m_q \pm 1\rangle \otimes |j, m\rangle)$$
$$\quad + \hbar\sqrt{(j \mp m)(j \pm m + 1)} \langle j', m' | (|k, m_q\rangle \otimes |j, m \pm 1\rangle) \qquad (15.128)$$

Thus, we find the same recursion relation connecting
$\langle \alpha', j', m' - 1 | O_q^{(k)} | \alpha, j, m \rangle$ to $\langle \alpha', j', m' | O_{q-1}^{(k)} | \alpha, j, m \rangle$ and
$\langle \alpha', j', m' | O_q^{(k)} | \alpha, j, m + 1 \rangle$ as the relation connecting
$\langle \alpha', j', m' - 1 | (|k, m_q\rangle \otimes |j, m\rangle)$ to $\langle \alpha', j', m' | (|k, m_q + 1\rangle \otimes |j, m\rangle)$ amd
$\langle \alpha', j', m' | (|k, m_q\rangle \otimes |j, m + 1\rangle)$ (and similarly for the sign-flipped version,
which is not important here).

Now, we make a subtle argument. The connection between matrix elements of different (m', m_q, m) must be independent of α and α' because the recursion relation is the same as that for Clebsch-Gordan coefficients, which do not depend on α and α'. As we have explained before in connection to the Clebsch-Gordan coefficients, once you set the initial conditions for the recursion by setting the matrix elements for $(m', m_q, m) = (j', k, j)$, then the remainder are all determined by the recursion relation. The value of the matrix element for $(m', m_q, m) = (j', k, j)$ cannot depend on (m', m_q, m) (but may depend on (j', k, j)). Thus, we are assured the matrix element can be factorized as stated earlier,

$$\langle \alpha', j', m' | O_q^{(k)} | \alpha, j, m \rangle = \langle \alpha', j' || \mathcal{O}^{(k)} || \alpha, j \rangle \langle j', m' | (|k, m_q\rangle \otimes |j, m\rangle) \tag{15.129}$$

where the double bars and lack of a q subscript on $\mathcal{O}^{(k)}$ in the first factor is meant to indicate that the matrix element will be independent of q, though in practice one must usually put in a specific $O_q^{(k)}$ to evaluate it.

Let's consider some examples of applications of the Wigner-Eckart theorem:

▶ **Projection Theorem**

Given the angular momentum vector operator $\vec{J}$ and some other arbitrary vector operator $\vec{V}$, one can show

$$\langle \alpha', j, m' | V_q | \alpha, j, m \rangle = \frac{\langle \alpha', j, m | \vec{J} \cdot \vec{V} | \alpha, j, m \rangle}{\hbar^2 j (j+1)} \langle j, m' | (|1, m_q\rangle \otimes |j, m\rangle) \tag{15.130}$$

What this does is let one calculate $\langle \alpha', j, m | \vec{J} \cdot \vec{V} | \alpha, j, m \rangle$ once for each (α, α', j) set and calculate $\langle j, m' | (|1, m_q\rangle \otimes |j, m\rangle)$ once for each (j, m, m') set and use those results for all combinations thereof. That is, regardless of how complicated V_q is, the angular part of it is trivial and the radial part only depends on the scalar product operators $\vec{J} \cdot \vec{V}$, which is general simpler that V_q.

Spherical Tensor Operators (cont.)

▶ **Relative Rates of Transitions**
Let's consider ratios of the magnitudes of matrix elements for the same
$(\alpha', \alpha, j', j, k)$ but different (m', m_q, m). The Wigner-Eckart theorem tells
us

$$\frac{\left|\langle \alpha', j', m_2' | O_{q_2}^{(k)} | \alpha, j, m \rangle\right|^2}{\left|\langle \alpha', j', m_1' | O_{q_1}^{(k)} | \alpha, j, m \rangle\right|^2} = \frac{\left|\langle j', m_2' | \left(|k, m_{q_2}\rangle \otimes |j, m\rangle\right)\rangle\right|^2}{\left|\langle j', m_1' | \left(|k, m_{q_1}\rangle \otimes |j, m\rangle\right)\rangle\right|^2} \quad (15.131)$$

That is, the ratios are determined entirely by CG coefficients.

Where might one use this? Suppose one has some system on which is
imposed a perturbing interaction of unknown form — *e.g.*, level transitions
in nuclei, decay modes of charmonium (atom formed from charm and
anticharm quarks). Transition rates between states are set by the
magnitude of the interaction's operator $O_q^{(k)}$ between those states. If we
use the above relation with $q_1 = q_2 = q$, then we see that the relative
rates of transitions to different final states from a given initial state is
determined only by the quantum numbers of the initial and final states and
the (k, m_q) values. Thus, if one measures the relative rates, one can
determine (k, m_q) because these ratios will be different for different
(k, m_q). That is, one can determine the nature of the interaction – scalar,
vector, tensor – without *any* detailed knowledge of its matrix elements.
This is *tremendously* powerful in complicated systems such as nuclei or
systems of fundamental particles.

Spherical Tensor Operators (cont.)

Connecting Cartesian and Spherical Tensor Operators

The obvious question to ask is – how do these objects relate to Cartesian tensor operators? We would guess that the relation is similar to that between Cartesian tensor states and spherical tensor states. But that discussion required the states to live in a Hilbert space, which allowed us to use the concepts of bases, direct products, and direct sums to understand how the spherical tensor states are a subspace of Cartesian tensor states. We do not (necessarily) have that structure for operators.

Fortunately, the demonstration that the action of $O_q^{(k)}$ is like adding the angular momentum $|k, m_q\rangle$ (*i.e.*, taking the direct product with this ket) lets us see there is a one-to-one mapping between spherical tensor operators and spherical tensor basis elements $\{|j, m\rangle\}$. Therefore, the proof of the relationship between Cartesian and spherical tensor states – that a Cartesian tensor space can be written as a direct sum of spherical tensor spaces – carries over to spherical tensor operators. Thus, the connection is the same: if the Cartesian tensor space $\tau^{(n)}$ has $C_j^{(n)}$ copies of the spherical tensor space $V^{(j)}$ in it, then the Cartesian tensor operator $O_{i_1 \cdots i_n}$ can be written as a sum of spherical tensor operators with $C_j^{(n)}$ spherical tensor operators of rank j contributing.

Applications of Spherical Tensor Operators for Understanding Degeneracies

Many problems involving spherical potentials have degeneracies in which states of different m or l have the same energy. We seek to explain these in terms of additional symmetries engendered by tensor operators.

The generic example is the lack of dependence of the energy on m for any spherically symmetric Hamiltonian; *i.e.*, Hamiltonians for which $[H, \vec{L}] = 0$. This symmetry of the Hamiltonian implies that the energy will be independent of the eigenvalue of L_z (or L_x or L_y, if either of those were chosen). This allows us to introduce the raising and lowering operators $L_\pm$ that commute with H and therefore allow one to change m value without changing E. These three steps are summarized as

1. Find a symmetry transformation of H
   Here: rotations about any axis.

2. Find the generator for the symmetry transformation.
   Here: $\vec{L}$

3. Construct a ladder operator from the generator to connect states that have different eigenvalues of L^2 and/or L_z but which are degenerate in energy. The ladder operator will of course commute with H because it is constructed from the generator of a symmetry transformation.

By the way, we note that the lack of dependence of the energy on the m eigenvalue is therefore a result of not just $[H, L_z] = 0$ and $[H, L^2] = 0$, but of $[H, \vec{L}] = 0$ because one would not have been able to construct the $L_\pm$ ladder operators if $[H, L_x] \neq 0$ or $[H, L_y] \neq 0$.

▶ **The Free Particle**
  The degeneracy is the fact that E depends on k alone, without any dependence on l or m. The additional symmetry transformation is translation, which is generated by $\vec{P}$. So we define raising and lowering operators

  $$P_\pm = \mp \frac{1}{\sqrt{2}} \left(P_x \pm i\, P_y \right) \tag{15.132}$$

  Since $\vec{P}$ is a vector operator, the above combinations $P_\pm$ are known to be the $m = \pm 1$ ($q = 1$ and 3) components of a spherical tensor of rank 1, and therefore acting on a state $|l, m\rangle$ with $P_\pm$ is like adding angular momentum $|1, \pm 1\rangle$.

So, now we act on the state $|k, l = 0, m = 0\rangle$ with P_+. Because $[H, \vec{P}] = 0$, $[H, P_+] = 0$ and so P_+ does not change the energy. Using our rules for angular momentum addition, we recognize that

$$(P_+)^l |k, 0, 0\rangle \propto |k, l, l\rangle \tag{15.133}$$

This angular momentum addition is unusually simple: if one acts with $(P_+)^l$, the m quantum number must be raised by l, so $m = l$ is necessary, and the maximum j state available after acting l times is $j = l$, and the only state with both $m = l$ that is accessible if the maximum allowed j is l is the state $|j = l, m = l\rangle$. So the ladder operator connects states of the same k and different l. Then, to get states of lower m for a given l, one uses the standard L_- operator, which of course commutes with H because $[H, \vec{L}] = 0$. Or, one could include applications of P_- to go directly from $|k, l, m\rangle$ to $|k, l + 1, m\rangle$ for $m < l$ (one will of course get mixtures of states of different l by doing this).

One may of course calculate explicitly the wavefunctions for the $|k, l, l\rangle$ states in this way by projecting onto the $\{|\vec{r}\rangle\}$ basis, which turns $|k, l, l\rangle$ into a wavefunction and P_+ into a differential operator. This is shown in Shankar, we will not repeat it here.

▶ **Isotropic 3D SHO**

Recall that, for the isotropic 3D SHO, the energy eigenvalues and allowed values of l are

$$E_{n,l,m} = \left(n + \frac{3}{2}\right)\hbar\omega \qquad l = n, n-2, \ldots, 1 \text{ or } 0 \tag{15.134}$$

Again, there is degeneracy in all l (and of course m) for the same n. Moreover, there is a restriction on the allowed values of l that they must be spaced by 2 for any given n.

The additional unrevealed symmetry in the 3D isotropic SHO that creates this degeracy among different l (as well as the "step-2" structure in the allowed l) is made evident by rewriting the SHO Hamiltonian in a different form. We introduce raising and lowering *vector* operators:

$$\vec{a} = \frac{1}{\sqrt{2\,\mu\,\omega\,\hbar}}\left(\mu\,\omega\,\vec{R} + i\,\vec{P}\right) \tag{15.135}$$

$$\vec{a}^{\dagger} = \frac{1}{\sqrt{2\,\mu\,\omega\,\hbar}}\left(\mu\,\omega\,\vec{R} - i\,\vec{P}\right) \tag{15.136}$$

The components of these vector operators are just the raising and lowering operator for the 1D SHO's in each coordinate axis. They therefore satisfy

$$[a_i, a_j^\dagger] = \delta_{ij} \tag{15.137}$$

The Hamiltonian may be rewritten

$$H = \left(\vec{a}^\dagger \cdot \vec{a} + \frac{3}{2} \right) \hbar \omega \tag{15.138}$$

The original Hamiltonian was invariant under orthogonal rotation transformations in real 3-dimensional space. The new Hamiltonian is manifestly invariant under unitary rotation transformations in *complex* 3-dimensional space, a 3-dimensional vector space in which the components are allowed to be complex numbers.

The expansion of the space expands the number of free parameters and hence generators that the symmetry transformations have. A real orthogonal matrix 3×3 matrix consists of 9 real numbers with 3 orthogonality conditions and 3 normalization conditions, so there are only three free parameters. This is reflected in the fact that sucha matrix always be written in the form $R = \exp(\vec{\theta} \cdot \vec{\mathbf{M}})$, with three free parameters and three generators. A complex unitary 3×3 matrix has 9 complex parameters (18 real numbers). The orthogonality conditions now result in loss of 6 parameters because each orthogonality condition has a real and a complex part. The normalization conditions still only impose 3 conditions (because the normalization conditions are automatically real equations). So there are 9 free parameters left, requiring 9 generators.

What are the nine generators? They are the quantities

$$O_{ij} = a_i^\dagger a_j \qquad (15.139)$$

because each one of these conserves the total energy of the system – one unit of energy is added, one is subtracted. The states are of course changed by O_{ij}, but the total energy is not. Clearly, this reflect degeneracy in the energy spectrum. Since $\vec{a}$ and $\vec{a}^\dagger$ are vector operators, their outer product is a tensor operator. In its current form, though, it is useless for telling us about how degenerate states with different angular momentum are connected because it is in Cartesian tensor form. We must rewrite it as a spherical tensor operator. These will be the necessary ladder operators because they add and subtract angular momentum.

Fortunately, our generic decomposition of Cartesian tensor operator of rank 2 into spherical tensor operators applies directly here. Unlike our example for the moment of inertia tensor, though, the components of the two vectors involved in the outer product are not the same and sometimes do not commute, so we must be careful to preserve the ordering. Simply applying those formulae, we obtain

$$O^{(2)}_{3\pm 2} = \frac{1}{2}\left[\left(a^\dagger_1 a_1 - a^\dagger_2 a_2\right) \pm i\left(a^\dagger_1 a_2 + a_1 a^\dagger_2\right)\right] \tag{15.140}$$

$$O^{(2)}_{3\pm 1} = \frac{1}{2}\left[\mp\left(a^\dagger_1 a_3 + a^\dagger_3 a_1\right) - i\left(a^\dagger_2 a_3 + a^\dagger_3 a_2\right)\right] \tag{15.141}$$

$$O^{(2)}_3 = -\sqrt{\frac{1}{6}}\left[a^\dagger_1 a_1 + a^\dagger_2 a_2 - a^\dagger_3 a_3\right] \tag{15.142}$$

$$O^{(1)}_{2\pm 1} = -\frac{1}{2}\left[\left(a^\dagger_1 a_3 - a^\dagger_3 a_1\right) \pm i\left(a^\dagger_2 a_3 - a^\dagger_3 a_2\right)\right] \tag{15.143}$$

$$O^{(1)}_2 = i\sqrt{\frac{1}{2}}\left[a^\dagger_1 a_2 - a^\dagger_2 a_1\right] \tag{15.144}$$

$$O^{(0)}_1 = -\sqrt{\frac{1}{3}}\left[a^\dagger_1 a_1 + a^\dagger_2 a_2 + a^\dagger_3 a_3\right] \tag{15.145}$$

Spherical Tensor Operators (cont.)

We may rewrite the rank 0 piece in a way that makes it clear it is a vector

$$O_1^{(0)} = -\sqrt{\frac{1}{3}}\, \vec{a}^\dagger \cdot \vec{a} \tag{15.146}$$

This object is clearly invariant under rotations in the space of complex 3D vectors; it is proportional to the Hamiltionian (with an offset), thus reflecting the fact that the Hamiltonian is invariant under such rotations.

And similarly for the rank 1 piece:

$$O_{2\pm 1}^{(1)} = -\frac{1}{2}\left[-\left(\vec{a}^\dagger \times \vec{a}\right)_2 \pm i\left(\vec{a}^\dagger \times \vec{a}\right)_1\right] \tag{15.147}$$

$$= \left(i\sqrt{\frac{1}{2}}\right)\left(\mp\sqrt{\frac{1}{2}}\left[\left(\vec{a}^\dagger \times \vec{a}\right)_1 \pm i\left(\vec{a}^\dagger \times \vec{a}\right)_2\right]\right) \tag{15.148}$$

$$O_2^{(1)} = \left(i\sqrt{\frac{1}{2}}\right)\left(\vec{a}^\dagger \times \vec{a}\right)_3 \tag{15.149}$$

These are the same relations between the components of a vector operator and its corresponding spherical tensor operator of rank 1 that we found before; so, the rank 1 piece is the Cartesian vector $i\left(\vec{a}^\dagger \times \vec{a}\right)/\sqrt{2}$.

Now, one ought to worry that the above vector operator can raise angular momentum by one unit, giving the state $|n, l + 1, l + 1\rangle$ from the state $|n, l, l\rangle$. Fortunately, the operator vanishes:

$$\vec{a}^\dagger \times \vec{a} \propto \left(a\vec{R} - i\,b\,\vec{P}\right) \times \left(a\vec{R} + i\,b\,\vec{P}\right) \tag{15.150}$$

$$= a^2 \vec{R} \times \vec{R} + b^2 \vec{P} \times \vec{P} + i\,a\,b\left(\vec{R} \times \vec{P} - \vec{P} \times \vec{R}\right) \tag{15.151}$$

$$= 2\,i\,a\,b\,\vec{L} = \frac{i}{\hbar}\vec{L} \tag{15.152}$$

where we used $\vec{P} \times \vec{R} = -\vec{R} \times \vec{P}$ because all products of components of $\vec{R}$ and $\vec{P}$ in the cross product commute (*e.g.*, $R_1 P_2$, etc.) We also inserted the values for a and b from the definition of $\vec{a}$. So, we see that $i\,(\vec{a}^\dagger \times \vec{a})\,/\sqrt{2} = -\vec{L}/\sqrt{2}$, so then

$$O^{(1)}_{2\pm 1} = \mp L_\pm / \sqrt{2} \qquad O^{(1)}_2 = -L_z / \sqrt{2} \tag{15.153}$$

Our worry was that $O^{(1)}_1 |n, l, l\rangle \propto |n, l + 1, l + 1\rangle$. But we know that $L_+|n, l, l\rangle = 0$. We in fact know the action of $L_\pm$ or L_z on any state $|n, l, m\rangle$ cannot change l, so we have no worry about connecting states of consecutive l.

One footnote on this: one might think that the Wigner-Eckart theorem tells us, because $\vec{L}$ itself is a spherical tensor of rank 1, that, for example, $L_-|n, l, l\rangle$ ought to have contributions from $|n, l-1, l-1\rangle$, $|n, l, l-1\rangle$, and $|n, l+1, l-1\rangle$ because the Clebsch-Gordan coefficients would indicate it. The thing that is missing in that argument is the fact that the matrix element of $\vec{L}$ between states of different l vanishes due to the details of the definition of $\vec{L}$. That is, though $\vec{L}$ is a spherical vector operator and thus is subject to the Wigner-Eckart theorem, that theorem says nothing about the value of $\langle \alpha', j' ||\mathcal{O}^{(k)}||\alpha, j\rangle$, which is perfectly well allowed to vanish for $j' \neq j$. So there is no contradiction with the Wigner-Eckart theorem.

The point of the above exercise is to see that the $\mathcal{O}^{(j)}$ spherical tensor operators are the necessary ladder operators to connect degenerate states of different l. The operator $O_1^{(2)}$ acting on a state $|n, l, l\rangle$ adds 2 to both l and m (remember, this is particular to acting on the highest m state for a given l with the highest m_q operator of a given rank k). So multiple applications of $O_1^{(2)}$ allows one to access the top m state for each value of l allowed for a given n, just as P_+ applied multiple times did the same for the free particle. One can then use the standard angular momentum lowering operators L_- to obtain states of lower m for each l. One could also reach these states of lower m by application of $O_q^{(2)}$ for $q > 1$ directly from $|n, 0, 0\rangle$ or $|n, 1, 1\rangle$.

Representation Theory for Angular Momentum

Introduction

The goal of this section is to give a brief overview of how the structure we have seen for angular momentum is tied to the underlying mathematical structure of rotations. This discussion will hopefully provide a somewhat unified picture of all that we have seen so far, in particular tying together spaces $V^{(j)}$ of different j and Cartesian and spherical tensor states and operators.

Definitions

First, we must make some definitions:

▶ **Groups**
A **group** is defined to be a set of elements with a binary operation rule that specifies how combinations of pairs of elements yield other members of the group. The set must be **closed** under the operation – binary combinations of members of the set may only yield other members of the set – to be called a group. In addition, the binary operation must be associative, $a(b\,c) = (a\,b)\,c$, there must be an identity element 1 such that $a\,1 = a$, and each element must have an inverse such that $a^{-1}\,a = a\,a^{-1} = 1$. Note that the group operation need not be commutative.

▶ **Fields**
A **field** is a group that has two kinds of operations, addition and multiplication. It is a group under each of the operations separately, and in addition satisfies distributivity: $a(b + c) = a\,b + a\,c$.

▶ **Group Algebras**
A **group algebra** is a combination of a field F (with addition $+$ and multiplication $\cdot$) and a group G (with multiplication $*$), consisting of all *finite* linear combinations of elements of G with coefficients from F, $a\,g + b\,h$, where a and b belong to F and g and h belong to G. The group operations on F and G continue to work:

$$
\begin{aligned}
a\,g + b\,g &= (a+b)\,g \\
a \cdot \sum_i a_i\,g_i &= \sum_i (a \cdot a_i)\,g_i \\
\left(\sum_i a_i\,g_i \right) * \left(\sum_j b_j\,h_j \right) &= \sum_{i,j} (a_i \cdot b_j)\,(g_i * h_j)
\end{aligned}
$$

▶ **Lie Algebras**
A **Lie algebra** is a group algebra with the additional conditions that the group elements $\{\sigma_i\}$ belonging to G satisfy the commutation relations

$$[\sigma_i, \sigma_j] \equiv \sigma_i \sigma_j - \sigma_j \sigma_i = c_{ij}^k \sigma_k$$

The $\{c_{ij}^k\}$ are called the **structure constants** of the Lie algebra. They satisfy $c_{ij}^k = -c_{ji}^k$. (Actually, the real definition of a Lie algebra is slightly more generic, allowing for definition of a commutator without G being part of a group algebra.) A Lie algebra may have a finite or infinite number of group members, though in general we will only consider finite ones. The matrices $\mathbf{M}_x$, $\mathbf{M}_y$, and $\mathbf{M}_z$ defined earlier form a Lie algebra with the real numbers as the field F.

Note that, even though we use σ_i for the algebra members above, they are not necessarily the Pauli matrices (we are running out of Greek letters...).

▶ **Lie Groups**
A **Lie group** is the exponential of the Lie algebra, consisting of all possible
elements

$$a = \exp\left(\sum_k \theta_k \sigma_k\right)$$

where the θ_k are members of the field F and the σ_k are members of the group
G. The exponential is defined in terms of its infinite power series expansion,
which is well-defined for a group. Note that the members of the Lie group are not
members of the Lie algebra because, by dint of the power series expansion, they
are *infinite* linear combinations of members of the Lie algebra. The Lie group is
entirely separate from the Lie algebra. The Lie group is, as its name indicates, a
group. Moreover, thanks to the method of definition, the group is differentiable
with respect to the members of the field. An example of a Lie group we have
encountered is $SO(3)$, the set of 3×3, orthogonal, determinant 1 matrices that
perform spatial rotations in three dimensions; its algebra is the Lie algebra
consisting of $\mathbf{M}_x$, $\mathbf{M}_y$, and $\mathbf{M}_z$ defined above.

Representations

Fundamentally, all that is needed to specify any of the above objects are multiplication and addition tables, the concept of convergence of an infinite series, and a rule for differentiation. The group elements need not be written out in terms of some simpler objects. Of course, in practice, we almost always do write the group elements out as matrices and rely on the rules for matrix multiplication, for multiplication and addition of complex numbers, for infinite series in complex numbers, and for differentiation of complex functions to specify all of the necessary group operations. Such a mapping of group elements to matrices is called a **representation**. A representation may be more general than that — *i.e.*, maybe the group can be mapped to some other kind of object on which the rules for all these operations are known — but matrix representations are the ones we will primarily deal with.

Representation Theory for Angular Momentum (cont.)

The Example of *SO*(3)

As we noted above, one Lie group that we have been spending a lot of our time on is *SO*(3), the group of "special" (determinant 1) orthogonal transformations on 3 spatial dimensions. This group is essentially defined in terms of its representation in terms of 3×3, orthogonal, determinant 1 matrices of real numbers, so it is difficult to separate in one's mind the group and this particular representation. The Lie algebra consists of the three matrices $\mathbf{M}_x$, $\mathbf{M}_y$, and $\mathbf{M}_z$ and their commutation relation

$$[\mathbf{M}_i, \mathbf{M}_j] = \epsilon_{ijk}\mathbf{M}_k \tag{15.154}$$

The field consists of real numbers. Addition and multiplication in the Lie algebra is specified completely by the rules for matrix multiplication of 3×3 matrices. The Lie group is defined by the power series expansion of the exponential. The convergence criterion and differentiation are simply defined in terms of element-by-element convergence and differentiation.

Other Representations of *SO*(3)

But we have encountered many other representations of *SO*(3) also. We may rewrite the Lie algebra in the form of the matrices $i\,\hbar\,\mathbf{M}_i$, $i = 1, 2, 3$, and the commutation relation

$$[i\,\hbar\,\mathbf{M}_i, i\,\hbar\,\mathbf{M}_j] = i\,\hbar\,\epsilon_{ijk}\,(i\,\hbar\,\mathbf{M}_k) \tag{15.155}$$

The structure constants are different, so this is a different Lie algebra. But we can generate the same Lie group by considering a different field:

$$a = \exp\left(-\frac{i}{\hbar}\sum_k \theta_k\,(i\,\hbar\,\mathbf{M}_k)\right) \tag{15.156}$$

These yield the same real, orthogonal, determinant 1, 3×3 matrices that the original Lie algebra yielded, so they are the same Lie group.

Using operator methods, we showed that there are an infinite number of representations of this alternate version of the Lie algebra; the angular momentum j representation consists of the three $(2j+1) \times (2j+1)$ matrices $\mathbf{J}_i^{(j)}$, $i = 1, 2, 3$, with

$$\left(\mathbf{J}_1^{(j)}\right)_{pq} = \frac{\hbar}{2} \left[\delta_{m_p, m_q+1} \sqrt{(j - m_q)(j + m_q + 1)} + \delta_{m_p, m_q-1} \sqrt{(j + m_q)(j - m_q + 1)} \right]$$
(15.157)

$$\left(\mathbf{J}_2^{(j)}\right)_{pq} = \frac{\hbar}{2i} \left[\delta_{m_p, m_q+1} \sqrt{(j - m_q)(j + m_q + 1)} - \delta_{m_p, m_q-1} \sqrt{(j + m_q)(j - m_q + 1)} \right]$$
(15.158)

$$\left(\mathbf{J}_3^{(j)}\right)_{pq} = \hbar \, m_q \, \delta_{m_p m_q}$$
(15.159)

$$m_p = j - p + 1 \qquad m_q = j - q + 1 \qquad m_p, m_q = -j, -j+1, \ldots, j$$
(15.160)

with j integer or half-integer.

A Lie group formed by the above kind of exponentials

$$a = \exp\left(-\frac{i}{\hbar}\sum_k \theta_k \left(\mathbf{J}_k^{j)}\right)\right) \tag{15.161}$$

must therefore also be a representation of $SO(3)$. But we see that these representations look very different, as they consist of $(2j+1) \times (2j+1)$ matrices, possibly including complex numbers. Even the $j = 1$ representation looks different than the Lie group generated by the $i\,\hbar\,\mathbf{M}_i$ matrices because the $\mathbf{J}_i^{(1)}$ matrices are written in a basis consisting of the eigenvectors of $i\,\hbar\,\mathbf{M}_3$, rather than the Cartesian basis that the $i\,\hbar\,\mathbf{M}_i$ are written in terms of.

Confused about *SO*(3) Representations?

There are thus two things that are terribly confusing about all of this. The first is the difference between the $i\,\hbar\,\mathbf{M}_i$ representation and the $\mathbf{J}_i^{(1)}$ representation (the $j = 1$ representation). That is just a matter of change of basis for the unit vectors of the space that the matrix representation acts on. (There is the minor complication that the Cartesian basis is completely real while the $j = 1$ basis is not, but one can recover any purely real vector by the appropriate linear combination of the $j = 1$ basis elements with complex coefficients).

The second and more confusing thing is that, for the j representation, using the $\mathbf{J}_i^{(j)}$ matrices, both the Lie algebra and the Lie group consist of $(2\,j + 1) \times (2\,j + 1)$ matrices, with possibly complex coefficients. They act on (possibly complex) column matrices with $2\,j + 1$ elements. So, how can these be rotations of vectors in real 3-dimensional space?

The key point is that the algebra and the group, once their addition, multiplication, and exponentiation tables are defined, no longer need to act on the objects they may have originally been intended for. These tables map pairs of algebra or group elements to the algebra or group, or map a set of field elements and the algebra elements to the group elements. The objects that the matrices act on enter nowhere. It might have been better to call the group by some other name and then note that its representation using the $i\,\hbar\,\mathbf{M}_i$ algebra is the set of rotation matrices in three spatial dimensions.

Representation Theory for Angular Momentum (cont.)

Representations and Irreducibility for $SO(3)$

The issue of irreducibility also has meaning in terms of representations. Consider the rotation operations on Cartesian tensors of rank n. These operations look complicated because they involve multiple copies of the rotation matrices. But recall that we showed that any Cartesian tensor of rank n can be written as a direct product of Cartesian tensors of rank 1 (vectors) and that a rotation operator acting on the direct product space is a direct product of rotation operators acting on the factor spaces. That is, if we write a Cartesian tensor of rank n as a sum of direct products of unit vectors:

$$\mathcal{T} = \sum_{i_1 \cdots i_n = 1}^{N} C_{i_1 \cdots i_n} \vec{e}_{i_1} \otimes \cdots \otimes \vec{e}_{i_n} \tag{15.162}$$

then a rotation operation on the tensor is

$$R^{(n)}_{\vec{\theta}} \mathcal{T} = \sum_{i_1 \cdots i_n = 1}^{N} C_{i_1 \cdots i_n} \left(\mathbf{R}_{\vec{\theta}} \vec{e}_{i_1} \right) \otimes \cdots \otimes \left(\mathbf{R}_{\vec{\theta}} \vec{e}_{i_n} \right) \tag{15.163}$$

That is, the n rotation matrices are tied together — they must have the same rotation angle $\vec{\theta}$ — so there is a one-to-one mapping between the "standard" $SO(3)$ matrix representation that consists of 3×3 matrices acting on 3-element column vectors and the rotation operators that act on Cartesian tensors of rank n. So, in spite of their very different apparent structure, the group of rotation operators acting on Cartesian tensors of rank n is also a representation of $SO(3)$.

But, we can show that this representation of $SO(3)$ is reducible as follows.

1. Because Cartesian and spherical tensors of rank 1 are the same thing, we are looking at rotation operators acting on direct products of n spherical tensors of rank 1.

2. The direct product states can be written as direct sums of states of spherical tensor states of ranks from 0 to n via our usual addition of angular momentum rules (*i.e.*, Clebsch-Gordan coefficients). Remember that this rewriting is just a unitary transformation of basis: if you order the direct product states and the direct sum states, and write any state as a column vector with N^n elements in either direct product basis or the direct sum basis, then the Clebsch-Gordan coefficients are the elements of a matrix that transforms between the two bases.

3. But, of course, the rotation operators acting on Cartesian tensors of rank n can then be represented as $N^n \times N^n$ matrices acting on the column vectors in the direct product basis. One can then apply the unitary transformation matrix to go from the direct product basis to the direct sum basis to get a matrix representation of the same rotation operators in the direct sum basis.

4. But, the group of rotation operators, even for Cartesian tensors of rank n, is just $SO(3)$ as we explained above. We already have matrix representations of $SO(3)$ in the $V^{(j)}$ subspaces of the direct sum space.

5. So, the matrix representation of the rotation operators in the direct sum basis will just look like block diagonal matrices with each block being a matrix representation of $SO(3)$ for the corresponding $V^{(j)}$ subspace.

Since the matrix representation is block diagonal, it fits the definition of reducibility – there are subspaces of the direct sum space that are invariant under the action of this matrix representation, so the matrix representation is reducible into blocks that act on each subspace independently.

Connection between $SO(3)$ and $SU(2)$

We make a final interesting point, concerning the relation between $SO(3)$ and another group $SU(2)$. There is a group called $SU(2)$ that is defined in terms of its representation as the set of all unitary, determinant 1, 2×2 matrices. We note that $SU(2)$ is obtained by considering the $j = 1/2$ representation of $SO(3)$ and allowing the rotation angle θ to have magnitude belonging to the interval $[2\,\pi, 4\,\pi)$. Recall that the rotation matrices for $j = 1/2$ can be written in closed form as

$$T^{(1/2)}(\vec{\theta}) = \cos\frac{\theta}{2} + 2\,\frac{i}{\hbar}\,\widehat{\theta}\cdot\vec{J}^{(1/2)}\,\sin\frac{\theta}{2} \tag{15.164}$$

You can show that, if you write out the above formula completely, then you have all the necessary freedom to obtain all matrices in $SU(2)$ if you allow the angle θ to be in the range $[0, 4\,\pi)$. But we are interested in how $SU(2)$ connects to $SO(3)$. For simplicity, consider $\vec{\theta} = \theta\,\widehat{z}$. We have then

$$T^{(1/2)}(\theta\,\widehat{z}) = \cos\frac{\theta}{2} + i\left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right]\sin\frac{\theta}{2} = \left[\begin{array}{cc} e^{i\,\theta/2} & 0 \\ 0 & e^{-i\,\theta/2} \end{array}\right] \tag{15.165}$$

We normally only consider θ in the range $[0, 2\pi)$. But consider $T^{(1/2)}((\theta + 2\pi)\,\widehat{z})$:

$$T^{(1/2)}((\theta + 2\pi)\,\widehat{z}) = - \left[\begin{array}{cc} e^{i\,\theta/2} & 0 \\ 0 & e^{-i\,\theta/2} \end{array} \right] \tag{15.166}$$

These differ from the elements in the $[0, 2\pi)$ range by an overall negative sign. We know the $[0, 2\pi)$ piece is a representation of $SO(3)$ because there is a well-defined one-to-one mapping from $SO(3)$ to the $[0, 2\pi)$ piece, as defined by the above equation for $T^{(1/2)}$. Therefore, there is a two-to-one mapping of the full $SU(2)$ group, which is the full $[0, 4\pi)$ range for θ, to $SO(3)$. The negative sign that is added in for the $[2\pi, 4\pi)$ piece does not affect probabilities — *i.e.*, the probability remains single-valued even though the state picks up a minus sign after rotation by 2π — so it is ok that this negative sign appeaars. The matrix representation of $SU(2)$ is called an unfaithful representation of $SO(3)$ because of the two-to-oneness of the mapping.

Section 16
Time-Independent Perturbation Theory

Lecture 43:
Time-Independent Perturbation Theory

Revision Date: 2008/03/03

Derivation

Overview

Suppose we have a Hamiltonian that is composed of two pieces, $H = H^{(0)} + \delta\, H^{(1)}$, where we have solved the eigenvalue-eigenvector problem for $H^{(0)}$, yielding energies $E_n^{(0)}$ and eigenstates $|n^{(0)}\rangle$, where $H^{(1)}$ is of the same size as $H^{(0)}$ (in a manner that will be quantified later) and $\delta \ll 1$, and where we want to find approximate solutions to the eigenvalue-eigenvector equation for H. $H^{(0)}$ is known as the unperturbed Hamiltonian and $\delta\, H^{(1)}$ as the perturbation or perturbing Hamiltonian.

The method we will use to find such an approximate solution is to assume, because $\delta \ll 1$, that the eigenvalues of H will deviate from those of $H^{(0)}$ by a power series in δ, and that we may obtain approximate eigenstates of H via an expansion in terms of the eigenstates $|n^{(0)}\rangle$ of $H^{(0)}$ with coefficients in powers of δ, and that we may decide the quality of the approximation by restricting ourselves to include terms of only up to a certain power in δ.

We will find such an expansion and check that our assumptions are valid. The expansion will be termed a perturbation series.

Derivation (cont.)

Formalism

In the above vein, let's assume that the eigenstates and eigenvalues of H have the above form:

$$|n\rangle = |n^{(0)}\rangle + |n^{(1)}\rangle + |n^{(2)}\rangle + \cdots \tag{16.1}$$

$$E_n = E_n^{(0)} + E_n^{(1)} + E_n^{(2)} + \cdots \tag{16.2}$$

where the term with superscript $^{(j)}$ is of order δ^j. We may assume $\langle n^{(0)}|n^{(j)}\rangle = 0$ for $j \neq 0$ because any contribution to $|n\rangle$ that is parallel to $|n^{(0)}\rangle$ would imply a piece of $\delta H^{(1)}$ that is proportional to $H^{(0)}$; any such contribution to $\delta H^{(1)}$ could be shifted into $H^{(0)}$. Then the eigenvalue-eigenvector equation for H and $|n\rangle$ is

$$H|n\rangle = E_n|n\rangle \tag{16.3}$$

$$\left(H^{(0)} + \delta H^{(1)}\right)\left(|n^{(0)}\rangle + |n^{(1)}\rangle + \cdots\right) = \left(E_n^{(0)} + E_n^{(1)} + \cdots\right)\left(|n^{(0)}\rangle + |n^{(1)}\rangle + \cdots\right) \tag{16.4}$$

$$\delta H^{(1)}|n^{(0)}\rangle + H^{(0)}|n^{(1)}\rangle + \cdots = E_n^{(1)}|n^{(0)}\rangle + E_n^{(0)}|n^{(1)}\rangle + \cdots \tag{16.5}$$

we have canceled the zeroth-order term $H^{(0)}|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle$. Additional terms are dropped because they are of order δ^2 or higher, by assumption (which must be checked in the end).

Next, since we know $\langle n^{(0)} | n^{(j)} \rangle = 0$ for $j \neq 0$, we can simplify the above equation by acting from the left with $\langle n^{(0)} |$:

$$\langle n^{(0)} | \delta H^{(1)} | n^{(0)} \rangle + \langle n^{(0)} | H^{(0)} | n^{(1)} \rangle = E_n^{(1)} \langle n^{(0)} | n^{(0)} \rangle + E_n^{(0)} \langle n^{(0)} | n^{(1)} \rangle \qquad (16.6)$$

$$E_n^{(1)} = \langle n^{(0)} | \delta H^{(1)} | n^{(0)} \rangle \qquad (16.7)$$

We used the fact that, to $\mathcal{O}(\delta)$, $\langle n^{(0)} | n^{(0)} \rangle = 1$ because $\langle n^{(0)} | n^{(j)} \rangle = 0$ for $j \neq 0$. The second term on each side vanishes because

$$\langle n^{(0)} | H^{(0)} | n^{(1)} \rangle = E_n^{(0)} \langle n^{(0)} | n^{(1)} \rangle = 0 \qquad (16.8)$$

The above expression for $E_n^{(1)}$ is thus our first-order energy correction. We see that our approximation $E_n^{(1)} = \mathcal{O}(\delta)$ holds as long as it holds that $\langle n^{(0)} | H^{(1)} | n^{(0)} \rangle = \mathcal{O}\left(E_n^{(0)}\right)$.

Derivation (cont.)

Now, let's go back and try to find an expression for $|n^{(1)}\rangle$, the term of $\mathcal{O}(\delta)$ in the expansion of $|n\rangle$. Let us now act on the eigenvector-eigenvalue equation to $\mathcal{O}(\delta)$ from the left with a different eigenstate of the unperturbed Hamiltonian, $\langle m^{(0)}|$ with $m \neq n$, because it will let us select out different terms in the $\mathcal{O}(\delta)$ eigenvector-eigenvalue equation:

$$\langle m^{(0)}|\delta H^{(1)}|n^{(0)}\rangle + \langle m^{(0)}|H^{(0)}|n^{(1)}\rangle = E_n^{(1)}\langle m^{(0)}|n^{(0)}\rangle + E_n^{(0)}\langle m^{(0)}|n^{(1)}\rangle \quad (16.9)$$

$$\langle m^{(0)}|\delta H^{(1)}|n^{(0)}\rangle + E_m^{(0)}\langle m^{(0)}|n^{(1)}\rangle = E_n^{(0)}\langle m^{(0)}|n^{(1)}\rangle \quad (16.10)$$

$$\langle m^{(0)}|n^{(1)}\rangle = \frac{\langle m^{(0)}|\delta H^{(1)}|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}} \quad (16.11)$$

We thus now have the coefficient of $|m^{(0)}\rangle$ in an expansion of $|n^{(1)}\rangle$ in terms of unperturbed eigenstates. Two conditions must be met for these terms to be $\mathcal{O}(\delta)$; we require $\langle m^{(0)}|H^{(1)}|n^{(0)}\rangle = \mathcal{O}\left(E_n^{(0)}\right)$ and $E_n^{(0)} - E_m^{(0)} = \mathcal{O}\left(E_n^{(0)}\right)$; the former can be violated if $H^{(1)}$ is too big (just as it can be violated for $E_n^{(1)}$) and the latter can be violated if there is an energy level of the unperturbed Hamiltonian that is degenerate with or very close to $E_n^{(0)}$, $|E_n^{(0)} - E_m^{(0)}| \ll E_n^{(0)}$. In the latter case, difficulty arises when the energy difference is small because $|n\rangle$ acquires large contributions from nearby states; in such cases, one must diagonalize the Hamiltonian including the perturbation to get accurate results.

Derivation (cont.)

So, we have

$$|n^{(1)}\rangle = \sum_{m \neq n} |m^{(0)}\rangle \frac{\langle m^{(0)} | \delta H^{(1)} | n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}} \tag{16.12}$$

We have to check the normalization of $|n^{(1)}\rangle$:

$$\langle n | n \rangle = \langle n^{(0)} | n^{(0)}\rangle + \mathcal{O}(\delta)\langle m^{(0)} | n^{(0)}\rangle + c.c. + \mathcal{O}(\delta^2)\langle m^{(0)} | m^{(0)}\rangle \tag{16.13}$$

$$= 1 + \mathcal{O}(\delta)^2 \tag{16.14}$$

where the term $\mathcal{O}(\delta)$ vanishes because $\langle m^{(0)} | n^{(0)}\rangle = 0$. Thus, to correct the normalization of $|n\rangle$, we would multiply by a factor

$$\left(1 + \mathcal{O}(\delta^2)\right)^{-1/2} = 1 - \frac{1}{2}\,\mathcal{O}(\delta^2) \tag{16.15}$$

This would give a $\mathcal{O}(\delta^2)$ correction to $|n\rangle$, which we may neglect. Therefore,

$$|n\rangle = |n^{(0)}\rangle + \sum_{m \neq n} |m^{(0)}\rangle \frac{\langle m^{(0)} | \delta H^{(1)} | n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}} + \mathcal{O}(\delta)^2 \tag{16.16}$$

Derivation (cont.)

Finally, there are frequently cases in which the first-order energy correction $E_n^{(1)}$ vanishes because of selection rules (as we shall see in the example below), so let's consider the second-order energy correction. To $\mathcal{O}(\delta^2)$, the eigenvector-eigenvalue equation is (after canceling the zeroth-order terms)

$$\delta H^{(1)}|n^{(0)}\rangle + H^{(0)}|n^{(1)}\rangle \delta H^{(1)}|n^{(1)}\rangle + H^{(0)}|n^{(2)}\rangle$$
$$= E_n^{(1)}|n^{(0)}\rangle + E_n^{(0)}|n^{(1)}\rangle + E_n^{(2)}|n^{(0)}\rangle + E_n^{(1)}|n^{(1)}\rangle \qquad (16.17)$$

Let's again act from the left with $\langle n^{(0)}|$ and cancel out the first-order terms using our formula for $E_n^{(1)}$:

$$\langle n^{(0)}|\delta H^{(1)}|n^{(1)}\rangle + \langle n^{(0)}|H^{(0)}|n^{(2)}\rangle$$
$$= \langle n^{(0)}|E_n^{(0)}|n^{(1)}\rangle + \langle n^{(0)}|E_n^{(2)}|n^{(0)}\rangle + \langle n^{(0)}|E_n^{(1)}|n^{(1)}\rangle \qquad (16.18)$$

The second term on the left side and the first and last terms on the right side vanish (even for $E_n^{(1)} \neq 0$) because $|n^{(0)}\rangle$ is an eigenstate of $H^{(0)}$ and is orthogonal to $|n^{(j)}\rangle$ for $j \neq 0$. This leaves

$$E_n^{(2)} = \langle n^{(0)}|\delta H^{(1)}|n^{(1)}\rangle \qquad (16.19)$$

Since $|n^{(1)}\rangle$ and $\delta H^{(1)}$ are $\mathcal{O}(\delta)$, this is indeed a $\mathcal{O}(\delta^2)$ correction to the energy.

The above can be rewritten in a useful form that lets us avoid calculating the first-order eigenstate shift explicitly if we have no need for it. Inserting our formula for $|n^{(1)}\rangle$ into our formula for $|E_n^{(2)}\rangle$, we have

$$E_n^{(2)} = \langle n^{(0)} | \delta H^{(1)} \sum_{m \neq n} |m^{(0)}\rangle \frac{\langle m^{(0)} | \delta H^{(1)} | n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \tag{16.20}$$

$$= \sum_{m \neq n} \frac{|\langle m^{(0)} | \delta H^{(1)} | n^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}} \tag{16.21}$$

This direct formula is in many cases easier to evaluate than calculating the state shift first and then plugging into the second-order energy shift formula, as we will see in connection with the hydrogen atom.

We develop a clever trick to evaluate the above sum. Let us rewrite it as

$$E_n^{(2)} = \sum_{m \neq n} \frac{\langle n^{(0)} \mid \delta H^{(1)} | m^{(0)} \rangle \langle m^{(0)} \mid \delta H^{(1)} | n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \tag{16.22}$$

Were it not for the energy denominator, we could use completeness to evaluate the above sum. But, suppose we can find an operator Ω such that

$$\delta H^{(1)} = \left[\Omega, H^{(0)}\right] \tag{16.23}$$

Derivation (cont.)

Then,

$$E_n^{(2)} = \sum_{m \neq n} \frac{\langle n^{(0)} \mid \delta H^{(1)} \mid m^{(0)} \rangle \langle m^{(0)} \mid [\Omega, H^{(0)}] \mid n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \tag{16.24}$$

$$= \sum_{m \neq n} \frac{\langle n^{(0)} \mid \delta H^{(1)} \mid m^{(0)} \rangle \langle m^{(0)} \mid \left(E_n^{(0)} - E_m^{(0)} \right) \Omega \mid n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \tag{16.25}$$

$$= \sum_{m} \langle n^{(0)} \mid \delta H^{(1)} \mid m^{(0)} \rangle \langle m^{(0)} \mid \Omega \mid n^{(0)} \rangle - \langle n^{(0)} \mid \delta H^{(1)} \mid n^{(0)} \rangle \langle n^{(0)} \mid \Omega \mid n^{(0)} \rangle \tag{16.26}$$

$$= \langle n^{(0)} \mid \delta H^{(1)} \Omega \mid n^{(0)} \rangle - \langle n^{(0)} \mid \delta H^{(1)} \mid n^{(0)} \rangle \langle n^{(0)} \mid \Omega \mid n^{(0)} \rangle \tag{16.27}$$

where we used completeness in the next to last step. If one is evaluating the second-order energy shift because the first-order energy shift vanishes, the second term will vanish because it is proportional to the first-order energy shift.

Time-Independent Perturbation Theory: Derivation

Of course, it is rather difficult to find such an operator Ω except in special cases. However, we note that, in order to have done the above trick, we only required

$$\delta H^{(1)}|n^{(0)}\rangle = \left[\Omega, H^{(0)}\right]|n^{(0)}\rangle \tag{16.28}$$

If this holds, then we would still be able to replace $\delta H^{(1)}|n^{(0)}\rangle$ with $\left[\Omega, H^{(0)}\right]|n^{(0)}\rangle$ as we did in the first step at the top of the page. Of course, one now has to find an operator Ω for every $|n^{(0)}\rangle$, but, since one is frequently only interested in the energy shift for the ground state or a few excited states, this may be sufficient.

Simple Applications

SHO + Electric Field

Shankar does a nice example in which one considers a perturbing Hamiltonian that is a constant electric field applied to the SHO (which must be charged, of course). Classically, we know that such a constant electric field just displaces the equilibrium position of the SHO but otherwise leaves it unaffected. So we know that there is also an exact QM solution with the energies shifted by $q E x_0$ where x_0 is the displaced equilibrium position. But one can treat the problem with perturbation theory, too, and that is done.

The example is interesting in three ways:

▶ The perturbing Hamiltonian can be written as a sum of (unperturbed) creation and annihilation operators, which provides an easy means to calculate the various matrix elements; this is clearly a nice general techique for any situation in which raising and lowering operators are available. Here, this is done by writing the perturbing X operator as $X = \sqrt{\hbar/2\,m\,\omega}\,(a + a^\dagger)$.

▶ Because of the above form for the perturbing Hamiltonian, the first-order energy shift $E_n^{(1)}$ vanishes, requiring one to calculate the second-order energy shift as we have done above.

▶ One sees that the perturbation series is essentially the Taylor series one would obtain by applying the translation operator to the unperturbed states.

Simple Applications (cont.)

Anharmonic SHO

For the sake of variety, let's consider an alternate application, a quartic anharmonic term added to the SHO potential:

$$H = H^{(0)} + \delta\, H^{(1)} = \frac{P^2}{2\,m} + \frac{1}{2}\, m\,\omega^2\, X^2 + \alpha\, X^4 \qquad (16.29)$$

This is an interesting case because the validity of perturbation theory depends not just on the parameter α but also on $\langle X^4 \rangle$ for the state that perturbation theory is being applied to. At the end, we will write down a formal criterion for perturbation theory to be valid for an energy level n, but it is pretty clear that the criterion will be something like "when the physical extent of the wavefunction gets too large, as it will as n increases, the energy correction will become too large."

Let's first write out the perturbing Hamiltonian in convenient form:

$$\delta H^{(1)} = \alpha X^4 = \frac{\alpha \hbar^2}{4 \, m^2 \, \omega^2} \left(a + a^\dagger \right)^4 \tag{16.30}$$

The first-order energy shift for SHO level n is therefore

$$E_n^{(1)} = \frac{\alpha \hbar^2}{4 \, m^2 \, \omega^2} \langle n^{(0)} \, | \left(a + a^\dagger \right)^4 | n^{(0)} \rangle \tag{16.31}$$

When the above expression is expanded out, there are 16 terms. However, we know that a term must have as many creation as annihilation operators in order for it not to vanish (one will end up with $\langle n^{(0)} \, | m^{(0)} \rangle$ for such terms) — a nice example of a selection rule. Thus, the only terms that contribute are

$$a^2 (a^\dagger)^2 \qquad a \, a^\dagger a \, a^\dagger \qquad a \, a^\dagger a^\dagger a \qquad (a^\dagger)^2 a^2 \qquad a^\dagger a \, a^\dagger a \qquad a^\dagger a \, a \, a^\dagger \tag{16.32}$$

The coefficients are indeed all unity, aside from the overall prefactor.

Recalling that $a^{\dagger}|n^{(0)}\rangle = \sqrt{n+1}|(n+1)^{(0)}\rangle$ and $a|n^{(0)}\rangle = \sqrt{n}|(n-1)^{(0)}\rangle$, we thus have (in the same order as the above listing)

$$E_n^{(1)} = \frac{\alpha\,\hbar^2}{4\,m^2\,\omega^2}\left((n+2)(n+1) + (n+1)^2 + (n+1)n + (n-1)n + n^2 + n(n+1)\right) \tag{16.33}$$

$$= \frac{3\,\alpha\,\hbar^2}{4\,m^2\,\omega^2}\,[2\,n\,(n+1) + 1] \tag{16.34}$$

Simple Applications (cont.)

To decide whether perturbation theory is valid, we need to evaluate whether $E_n^{(1)}/E_n^{(0)} \ll 1$. This ratio is

$$\frac{E_n^{(1)}}{E_n^{(0)}} = \frac{3\,\alpha}{2\,m^2\,\omega^3/\hbar}\,\frac{2\,n\,(n+1)+1}{2\,n+1} \tag{16.35}$$

So, perturbation theory is valid if

$$\frac{\alpha}{m^2\,\omega^3/\hbar} \ll \frac{2}{3}\,\frac{2\,n+1}{2\,n\,(n+1)+1} \tag{16.36}$$

We see that this is a *n*-dependent criterion, and that, because the right side gets smaller as *n* increases, the criterion may be satisfied for low *n* but not for high *n*. The expression reduces to a simpler form for $n = 0$,

$$\frac{\alpha}{m^2\,\omega^3/\hbar} \ll \frac{2}{3} \qquad \text{for } n = 0 \tag{16.37}$$

Clearly, if the criterion is not met for the ground state, it will not be satisfied for higher states. So this simplified criterion is a condition for perturbation theory to be valid at all.

It is instructive to rewrite the above in a more physical manner, in terms of $\langle X^4 \rangle_n$ and $\langle X^2 \rangle_n$. We have calculated $\langle X^4 \rangle_n$ above, and we know from our discussion of the SHO that

$$\langle X^2 \rangle_n = \frac{\hbar}{m\omega}\left(n + \frac{1}{2}\right) = \frac{1}{m\omega^2}E_n^{(0)} \qquad (16.38)$$

So, we have

$$\frac{E_n^{(1)}}{E_n^{(0)}} = \frac{\alpha \langle X^4 \rangle_n}{m\omega^2 \langle X^2 \rangle_n} \qquad (16.39)$$

So, it's simply a question of whether the mean anharmonic potential energy is small compared to the harmonic potential energy. Of course, the first version of the expression is more useful for explicit calculation, but this version tells us what that calculational criterion corresponds to, physically. This physical criterion makes good sense intuitively.

Simple Applications (cont.)

A variational estimate of the ground and excited state energies of this problem can also be obtained; a reasonable method is to use trial wavefunctions that look like the true SHO ground and excited states, but with the coefficient in the argument of the Gaussian allowed to vary to minimize the energy. Such a procedure yields, when the anharmonic term is small compared to the harmonic term,

$$E_0^{var} = \frac{\hbar\omega}{2} + \frac{3\,\alpha\,\hbar^2}{4\,m^2\omega^2} \qquad \psi_0^{var}(x) = A_0 \exp^{-\left(1 + \frac{3\,\alpha\,\hbar}{m^2\omega^3}\right)\frac{m\,\omega\,x^2}{2\,\hbar}} \tag{16.40}$$

$$E_1^{var} = \frac{3\,\hbar\omega}{2} + \frac{15\,\alpha\,\hbar^2}{4\,m^2\omega^2} \qquad \psi_1^{var}(x) = A_1\,x\,\exp^{-\left(1 - \frac{5\,\alpha\,\hbar}{m^2\omega^3}\right)\frac{m\,\omega\,x^2}{2\,\hbar}} \tag{16.41}$$

The values we obtain from the perturbation analysis are

$$E_0^{pert} = \frac{\hbar\omega}{2} + \frac{3\,\alpha\,\hbar^2}{4\,m^2\omega^2} \qquad E_1^{pert} = \frac{3\,\hbar\omega}{2} + \frac{15\,\alpha\,\hbar^2}{4\,m^2\omega^2} \tag{16.42}$$

That is, the estimates are identical! Thus, we see that our variational estimates are remarkably well-corroborated by this more systematic procedure using perturbation theory. It is, however, worth remembering that the variational technique can be applied in nonperturbative cases, where the Hamiltonian is not "close to" a Hamiltonian whose eigenvalues and eigenstates we know.

Let's think about the eigenstate correction. Now, we are interested in the matrix element:

$$\langle m^{(0)} | n^{(1)} \rangle = \frac{\langle m^{(0)} | \delta H^{(1)} | n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \tag{16.43}$$

The terms that we determined would contribute to the energy shift *do not* contribute here because they connect $|n^{(0)}\rangle$ to $|n^{(0)}\rangle$. Let's consider the other terms:

▶ There will be terms that have four powers of either a or $a^\dagger$. These connect n to $m = n \pm 4$.

▶ There will be terms that have three powers of $a^\dagger$ and one power of a and vice versa. These connect n to $m = n \pm 2$.

▶ The remaining terms will have two powers each of $a^\dagger$ and a, which we have already considered in calculating the energy shift.

We can understand why only the $\Delta n = 2$ and $\Delta n = 4$ states contribute in terms of the position-space wavefunctions if we imagine calculating the above matrix elements in that basis.

▶ In the integral, X^4 is replaced by the even function x^4. The integral for the matrix element will therefore only be nonzero if the two states involved have the same parity so the overall integrand is even. Since the parity of a SHO state is $(-1)^n$, states separated by an even Δn have the same parity and those separated by an odd Δn have opposite parity, so the coupling only to states with Δn even makes sense.

▶ The coupling to only "nearby" states — those with $\Delta n \leq 4$ — is determined by the fact that the position-basis wavefunction $\psi_n(x) = \langle x | n \rangle$ is a polynomial of order n in x times a Gaussian. The function $x^4 \, \psi_n(x)$ will then be orthogonal to any polynomial of order greater than $n + 4$ or less than $n - 4$.

Even with this restricted set of contributions, it is still not particularly instructive to write the corrections out. It is interesting, however, to point out that adding in contributions from these other states, without changing the coefficient in the argument of the multiplying Gaussian, will never make the resulting state look like a Gaussian with an adjusted coefficient in the argument as we obtained from the variational technique. Thus, we are reminded of the fact that, for the variational technique, the error in the energy estimate is only quadratically dependent on the error in the wavefunction — you do not need a very accurate wavefunction to get a very accurate energy estimate.

Selection Rules

Selection Rules

We saw in our examples above how one can demonstrate that various matrix elements of $H^{(1)}$ involved in the perturbation series calculations will vanish using simple properties of the states and the Hamiltonian, without full calculation. We can codify these rules in a simple theorem:

Consider an operator Ω that commutes with the perturbing Hamiltonian, $[\Omega, H^{(1)}] = 0$. Then, if $|\alpha_i, \omega_i\rangle$, $i = 1, 2$, denote common eigenstates of $H^{(0)}$ (not $H^{(1)}$!) and Ω, then it holds:

$$\langle \alpha_2, \omega_2 | H^{(1)} | \alpha_1, \omega_1 \rangle = 0 \qquad \text{unless} \quad \omega_1 = \omega_2 \tag{16.44}$$

The proof is easy:

$$0 = \langle \alpha_2, \omega_2 | [\Omega, H^{(1)}] | \alpha_1, \omega_1 \rangle \tag{16.45}$$

$$= \langle \alpha_2, \omega_2 | \left(\Omega H^{(1)} - H^{(1)} \Omega \right) | \alpha_1, \omega_1 \rangle \tag{16.46}$$

$$= (\omega_2 - \omega_1) \langle \alpha_2, \omega_2 | H^{(1)} | \alpha_1, \omega_1 \rangle \tag{16.47}$$

Heuristically, the idea is that if $H^{(1)}$ "carries no Ω charge" — which is the same as commuting with Ω — then it will not change the the ω eigenvalue and so can only connect states with the same ω eigenvalue. We can see this as follows:

$$\omega_1 \left(H^{(1)} |\alpha_1, \omega_1 \rangle \right) = H^{(1)} \omega_1 |\alpha_1, \omega_1 \rangle = H^{(1)} \Omega |\alpha_1, \omega_1 \rangle = \Omega \left(H^{(1)} |\alpha_1, \omega_1 \rangle \right) \quad \text{(16.48)}$$

That is, $H^{(1)} |\alpha_1, \omega_1 \rangle$ is an eigenstate of Ω with eigenvalue ω_1 if $[\Omega, H^{(1)}] = 0$, so the matrix element of $H^{(1)} |\alpha_1, \omega_1 \rangle$ with any other eigenstate of Ω will vanish (modulo existence of degenerate eigenstates).

We can extend the above rule as follows. Suppose $H^{(1)}$ does not commute with Ω. Then eigenstates of Ω are not eigenstates of $H^{(1)}$ and, in general, $H^{(1)} |\alpha_1, \omega_1 \rangle$ will not be the ω_1 eigenstate of Ω and may not be any eigenstate of Ω. Clearly, though, in order for the matrix element under consideration to be nonvanishing, the state $H^{(1)} |\alpha_1, \omega_1 \rangle$ must have some projection onto $|\alpha_2, \omega_2 \rangle$. Heuristically, we can say that $H^{(1)}$ must "carry some Ω charge" and that $H^{(1)}$ must "add enough Ω charge to ω_1 to obtain ω_2." These are very unrigorous statements and require some concrete examples for explanation.

Selection Rules (cont.)

Parity Selection Rules

Consider the parity operator, Π, which is both unitary and Hermitian. In many cases, $H^{(0)}$ commutes with Π and so eigenstates of $H^{(0)}$ are eigenstates of Π (have definite parity) — the SHO is an example. Then our above rules tell us two things.

First, if the perturbing Hamiltonian commutes with parity, then a matrix element must be between states of the same parity in order to be nonvanishing. Our anharmonic term commutes with parity (because $\Pi^{\dagger} X^4 \Pi = X^4$), so its matrix elements are only nonvanishing for states of the same parity. This explains why, for example, $\langle m^{(0)} | X^4 | n^{(0)} \rangle = 0$ for $|m - n|$ odd: all even m, n correspond to parity $+1$ states and all odd m, n correspond to parity -1 states, so X^4 can only connect two states of the same parity; *i.e.*, $|m - n|$ even.

Second, if the perturbing Hamiltonian does not commute with parity, then its matrix elements between states of the same parity will in general vanish and its matrix elements between states that differ in parity by the parity that the operator contributes will not vanish. This explains why the first-order energy shift for the SHO + electric field vanishes: it is the matrix element $\langle n^{(0)} | X | n^{(0)} \rangle$; since we have the same state on both sides, the operator in the middle must commute with parity to yield a nonvanishing matrix element, and it does not, $\Pi^{\dagger} X \Pi = -X$. It also implies that the first-order eigenstate correction $|n^{(1)} \rangle$ will contain contributions from $|m^{(0)} \rangle$ where $|m - n|$ is odd, so that the $|n^{(0)} \rangle$ and $|m^{(0)} \rangle$ have opposite parity and $\langle m^{(0)} | X | n^{(0)} \rangle \neq 0$.

Angular Momentum Selection Rules

Let's see how a more complicated case works, wherein we calculate the matrix elements of a component of a vector operator between two angular momentum eigenstates. Here, the operator "carries angular momentum," so we shall see that we can obtain selection rules based on our rules for angular momentum addition. In addition, the Wigner-Eckart theorem will simplify the calculation of the nonvanishing matrix elements.

Let's consider the matrix elements of a vector operator $\vec{V}$,

$$\langle \alpha_2, j_2, m_2 \, | \, \vec{V} \, | \alpha_1, j_1, m_1 \, \rangle \tag{16.49}$$

We know that we can write the components of the Cartesian vector operator $\vec{V}$ as linear combinations of components of the corresponding spherical vector operator $\mathcal{V}^{(1)}$:

$$V_1^{(1)} = -\frac{V_x + i \, V_y}{\sqrt{2}} \qquad V_3^{(1)} = \frac{V_x - i \, V_y}{\sqrt{2}} \qquad V_2^{(1)} = V_z \tag{16.50}$$

$$\Longleftrightarrow \qquad V_x = -\frac{V_1^{(1)} - V_3^{(1)}}{\sqrt{2}} \qquad V_y = i \, \frac{V_1^{(1)} + V_3^{(1)}}{\sqrt{2}} \qquad V_z = V_2^{(1)} \tag{16.51}$$

Therefore, the matrix elements of the components of the spherical tensor operator will give us the matrix elements of the components of the Cartesian vector operator.

In calculating the spherical vector matrix elements, we can derive and make use of selection rules. The key observation is that, for any spherical tensor operator of rank k, $\mathcal{O}^{(k)}$, we have shown (Section 15.3, *Action on $|j, m\rangle$ states*)

$$O_q^{(k)}|j, m\rangle \propto (|k, m_q\rangle \otimes |j, m\rangle) \tag{16.52}$$

Heuristically, we say the the operator $O_q^{(k)}$ "carries angular momentum $|k, m_q\rangle$ and adds it to the state it acts on;" the matrix element is nonvanishing only if the projection of the "angular-momentum-summed state" $|k, m_q\rangle \otimes |\alpha_1, j_1, m_1\rangle$ onto $|\alpha_2, j_2, m_2\rangle$ is nonvanishing. Based on the values j_2, m_2, we can immediately determined that some matrix elements will always vanish.

Selection Rules (cont.)

Specifically:

$$\langle \alpha_2, j_2, m_2 \,|\, V_1^{(1)} |\alpha_1, j_1, m_1 \,\rangle \propto \langle \alpha_2, j_2, m_2 \,|\, (|1, 1\,\rangle \otimes |\alpha_1, j_1, m_1\,\rangle) = 0 \qquad (16.53)$$
$$\text{for} \quad |j_1 - j_2| > 1 \quad \text{and/or} \quad m_2 \neq m_1 + 1$$

$$\langle \alpha_2, j_2, m_2 \,|\, V_3^{(1)} |\alpha_1, j_1, m_1 \,\rangle \propto \langle \alpha_2, j_2, m_2 \,|\, (|1, -1\,\rangle \otimes |\alpha_1, j_1, m_1\,\rangle) = 0 \qquad (16.54)$$
$$\text{for} \quad |j_1 - j_2| > 1 \quad \text{and/or} \quad m_2 \neq m_1 - 1$$

$$\langle \alpha_2, j_2, m_2 \,|\, V_2^{(1)} |\alpha_1, j_1, m_1 \,\rangle \propto \langle \alpha_2, j_2, m_2 \,|\, (|1, 0\,\rangle \otimes |\alpha_1, j_1, m_1\,\rangle) = 0 \qquad (16.55)$$
$$\text{for} \quad |j_1 - j_2| > 1 \quad \text{and/or} \quad m_2 \neq m_1$$

As Shankar states, there may be additional parity selection rules that can be applied if j_1 and j_2 are orbital angular momentum with known parity eigenvalues. But the above gives the generic result, where we have shown that, given a particular $|j_1, m_1\,\rangle$, there are only nine $|j_2, m_2\,\rangle$ states out of an infinite set of them for which the matrix element does not vanish, three for each component of the spherical tensor operator $\mathcal{V}^{(1)}$. Once we have calculated these nine matrix elements, we can calculate all possible matrix elements of the Cartesian vector operator $\vec{V}$ from linear combinations thereof. The amount of calculation we need to do has been drastically reduced by the derivation of these selection rules.

Selection Rules (cont.)

Furthermore, the Wigner-Eckart theorem tells us that the matrix elements are just the Clebsch-Gordan coefficients up to an overall normalization:

$$\langle \alpha_2, j_2, m_2 | V_q^{(1)} | \alpha_1, j_1, m_1 \rangle = \langle \alpha_2, j_2 || V^{(1)} || \alpha_1, j_1 \rangle \langle j_2, m_2 | (|k, m_q \rangle \otimes |j_1, m_1 \rangle) \tag{16.56}$$

where we recall that "$\langle \alpha_2, j_2 || V^{(1)} || \alpha_1, j_1 \rangle$" means "calculate $\langle \alpha_2, j_2, m_2 | V_q^{(1)} | \alpha_1, j_1, m_1 \rangle$ for whichever q, m_1, m_2 set you like (is most convenient) and then use that same matrix element for all other q, m_1, m_2 combinations." So, we see that we only need to calculate *one* matrix element explicitly — and we get to pick the easiest one — and then we just need to look up Clebsch-Gordan coefficients to get *all the other* matrix elements! What a drastic decrease in work, from calculating an infinite number of matrix elements to calculating a single one!

Degenerate Perturbation Theory

The Problem and the Solution

Standard perturbation theory breaks down when the expansion parameter δ becomes comparable to or large compared to 1. This can of course happen if the first-order energy shift gets too large, but that just implies that the perturbing Hamiltonian is too large and perturbation theory is just not valid for that perturbation. A more generic and dangerous failure mode, though, is when the unperturbed energy levels are explicitly degenerate or just very close to each other, so that one or more coefficients in the first-order eigenstate correction become too big; then, the condition

$$\left| \frac{\langle m^{(0)} | \delta H^{(1)} | n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \right| \ll 1 \tag{16.57}$$

fails to be met. (Of course, the condition can fail to be met if the numerator is too large, also, but that is much more like the first-order energy-shift problem of $\delta H^{(1)}$ being too big.)

Degenerate Perturbation Theory (cont.)

Shankar gives a reasonably good explanation of this phenomenon; textbooks normally don't try to explain it, they just show how to deal with it. The basic idea is as follows. We have some degeneracy in $H^{(0)}$ in the default unperturbed basis $\{|n^{(0)}\rangle\}$; denote the degenerate subspace by $\mathcal{D}_n$ where $E_n^{(0)}$ is the unperturbed degenerate energy. In general, $\delta H^{(1)}$ breaks the degeneracy. (Let's assume for now it breaks the degeneracy fully, we'll explain below what happens when it does not.) Suppose one explicitly constructs the eigenstates of $H^{(0)} + \delta H^{(1)}$ (which one can always do by numerical methods). Those eigenstates will depend on δ in some way, and so one should be able to allow $\delta \to 0$ and recover unperturbed, degenerate eigenstates.

But, because of the degeneracy in $H^{(0)}$, the resulting unperturbed eigenstates, which we will denote by $\{|\overline{n}^{(0)}\rangle\}$, may not be the same as the $\{|n^{(0)}\rangle\}$: there is freedom because of the degeneracy. Thus, if we watch how the eigenstates change as we turn on δ, we will see them suddenly jump from $\{|n^{(0)}\rangle\}$ to $\{|\overline{n}^{(0)}\rangle\}$ for any infinitesimal deviation of δ from 0, and then evolve smoothly with δ from there on. This nondifferentiable dependence on δ is nonsense in perturbation theory and is intimately connected to the divergence in the first-order eigenstate correction. The clear way to solve the problem is to start out with the $\{|\overline{n}^{(0)}\rangle\}$ instead of the $\{|n^{(0)}\rangle\}$ and perform perturbation theory on the former. Then the nondifferentiable behavior will be removed.

Finding the $\{|\overline{n}^{(0)}\rangle\}$ corresponds to diagonalizing $\delta H^{(1)}$ in the degenerate subspace $\mathcal{D}_n$. If we do so, then, when we turn on $\delta H^{(1)}$, there will be no sudden changes in the $\mathcal{D}_n$ contribution to one of the $|\overline{n}^{(0)}\rangle$ — the matrix elements $\langle\overline{m}^{(0)}|\delta H^{(1)}|\overline{n}^{(0)}\rangle$ vanish for $|\overline{m}^{(0)}\rangle \neq |\overline{n}^{(0)}\rangle$, so there is no mixing of different directions in $\mathcal{D}_n$.

Degenerate Perturbation Theory (cont.)

One can see this calculationally too: if one finds these eigenvectors, then one will have a set of states for which the off-diagonal matrix elements of $\delta H^{(1)}$ vanish in $\mathcal{D}_n$:

$$\langle \overline{m}^{(0)} | \delta H^{(1)} | \overline{n}^{(0)} \rangle = 0 \tag{16.58}$$

after diagonalization for $|\overline{m}^{(0)} \rangle$ and $|\overline{n}^{(0)} \rangle$ in $\mathcal{D}_n$. This is an exact statement, true to any order in δ, as one can always exactly diagonalize any Hermitian matrix, and the portion of $\delta H^{(1)}$ in $\mathcal{D}_n$ is Hermitian. The vanishing of the energy denominator is also an exact statement. Hence, the two effects cancel and the first-order correction divergence is tamed. This calculational point of view is not very rigorous because one really needs to take some sort of limit and invoke l'Hôpital's rule to show that $0/0 \rightarrow 1$, but there is no limit to be taken — the vanishing of the numerator and denominator are exactly true for $\delta \neq 0$ also.

Now, there is an important subtlety to remember: just because we have diagonalized $\delta H^{(1)}$ in $\mathcal{D}_n$, it does not hold that we have fully diagonalized $\delta H^{(1)}$ in the whole Hilbert space, and so, as $\delta H^{(1)}$ is turned on, the diagonalized basis $\{|\overline{n}^{(0)} \rangle\}$ will in general acquire contributions from states outside of $\mathcal{D}_n$ that contribute to the first-order eigenstate correction and second-order energy corrections. (The first-order energy correction will not depend on states outside of $\mathcal{D}_n$ because it is just $\langle \overline{n}^{(0)} | \delta H^{(1)} | \overline{n}^{(0)} \rangle$) So, one still needs to do perturbation theory, one just does it on the diagonalized basis $\{|\overline{n}^{(0)} \rangle\}$.

Section 17
The Hydrogen Atom

Lecture 44:
The Unperturbed Hydrogen Atom

Revision Date: 2008/02/26

Summary of the Unperturbed Hydrogen Atom

The hydrogen atom is just the problem of two particles of equal and opposite charge bound by the Coulomb potential,

$$V(r) = -\frac{e^2}{r} \tag{17.1}$$

(Gaussian units) where one particle has charge e (the proton) and the other one has charge $-e$ (the electron). The solution to the problem is very much a straightforward application of what we have done so far, so we refer you to Shankar for all the details. We will summarize the results here and note interesting aspects of the derivation. We will work only on the bound states; free states of course exist but are of less interest.

Because the potential is spherically symmetric, the eigenstates of the Hamiltonian will be eigenstates of L^2 and L_z. In the position basis, the eigenstates must therefore be of the form

$$\langle \vec{r} \,|E, l, m \rangle = \psi_{E,l,m}(r, \theta, \phi) = R_{E,l} Y_l^m(\theta, \phi) \tag{17.2}$$

Recall that the radial function cannot depend on m because of the spherical symmetry of the problem. Inserting this form into the position-basis eigenvalue-eigenvector equation for the Hamiltonian yields the radial equation

$$\left\{ -\frac{\hbar^2}{2\mu} \left[\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \right) - \frac{l(l+1)}{r^2} \right] - \frac{e^2}{r} \right\} R_{E,l}(r) = E\, R_{E,l}(r) \tag{17.3}$$

Unperturbed Hydrogen Atom (cont.)

As we did for the general case (Section 14.9), we assume a solution of the form

$$R_{E,l}(r) = \frac{U_{E,l}(r)}{r} \tag{17.4}$$

which results in a simpler differential equation

$$\left[-\frac{\hbar^2}{2\,\mu} \frac{d^2}{dr^2} - \frac{e^2}{r} + \frac{l\,(l+1)\,\hbar^2}{2\,\mu\,r^2} \right] U_{E,l}(r) = E\,U_{E,l}(r) \tag{17.5}$$

If one assumes a power series solution, the three different powers of r in the equation will result in a three-term recursion relation, which is complicated. Recall that, for the general case, we obtained the asymptotic form for $r \to \infty$ using the WKB approximation, finding

$$U_{E,l}(r) \overset{r \to \infty}{\propto} r^{\,\mu\,e^2/\hbar^2\kappa}\,e^{-\kappa\,r} \qquad \kappa = \sqrt{-\frac{2\,\mu\,E}{\hbar^2}} \tag{17.6}$$

and we also found that, for $r \to 0$, the nonsingular solution satisfies

$$U_{E,l}(r) \overset{r \to 0}{\propto} r^{l+1} \tag{17.7}$$

We therefore change to the dimensionless variable

$$\rho = \frac{r}{r_0} \qquad r_0 = \frac{1}{\kappa} = \sqrt{-\frac{\hbar^2}{2\,\mu\,E}} \tag{17.8}$$

and assume a solution of the form

$$U_{E,l}(r) \propto e^{-\rho}\, v_{E,l}(\rho) \tag{17.9}$$

The differential equation becomes

$$\frac{d^2}{d\rho^2}\, v_{E,l}(\rho) - 2\,\frac{d}{d\rho}\, v_{E,l}(\rho) + \left[\frac{e^2}{|E|\, r_0} \frac{1}{\rho} - \frac{l(l+1)}{\rho^2} \right] v_{E,l}(\rho) = 0 \tag{17.10}$$

We are left with an equation that will yield a simple two-term recursion relation. It will be simpler even than the SHO because the recursion relation will connect consecutive terms: we have only the powers ρ^{-1} and ρ^{-2} in the equation (note that a derivative $d/d\rho$ is like a power ρ^{-1}).

Owing to the known behavior as $r \to 0$, we assume a solution of the form

$$v_{E,l}(\rho) = \rho^{l+1} \sum_{k=0}^{\infty} C_k \rho^k \qquad (17.11)$$

The recursion relation is

$$C_{k+1} = \frac{-\frac{e^2}{|E| r_0} + 2 (k + l + 1)}{(k + l + 2) (k + l + 1) - l(l + 1)} \qquad (17.12)$$

As $k \to \infty$, this recursion relation asymptotes to

$$C_{k+1} \overset{k \to \infty}{\longrightarrow} \frac{2}{k} C_k \qquad (17.13)$$

which, if unterminated, results in a exponential with positive argument, yielding an unnormalizable wavefunction (recall, we ran into a simple problem with the SHO).

We therefore require termination of the power series, yielding the condition

$$\frac{e^2}{|E| \, r_0} = 2 \, (k + l + 1) \quad \Longleftrightarrow \quad E = -\frac{\mu \, e^4}{2 \, \hbar^2} \frac{1}{(k + l + 1)^2} \qquad k = 0, 1, 2, \ldots ; l = 0, 1, 2, \ldots \tag{17.14}$$

or, as it is more conventionally written

$$E_n = -\frac{\mu \, e^4}{2 \, \hbar^2} \frac{1}{n^2} \equiv -\frac{Ry}{n^2} \qquad n = 1, 2, \ldots ; l = 0, 1, \ldots, n-1 \tag{17.15}$$

where the *Rydberg*, Ry, sets the energy scale of the problem and n is known as the principal quantum number. The "accidental degeneracy," whereby the energy does not depend on l, is apparent; we shall understand it shortly in terms of a hidden symmetry whose generators yield spherical tensor ladder operators, similar to what we saw in Section 15.3. The degeneracy is

$$\sum_{l=0}^{n-1} (2 \, l + 1) = n^2 \tag{17.16}$$

The states are denoted by their $n \, l$ values, and we usually use the spectroscopic notation $s, p, d, f, g, h, \ldots$ for $l = 0, 1, 2, 3, 4, 5, \ldots$ so that states would be of the form $1s, 2s, 2p, 3s, 3p, 3d, \ldots$.

The full radial wavefunction is

$$R_{n,l}(r) = \exp\left(-\frac{r}{n\,a_0}\right)\left(\frac{r}{n\,a_0}\right)^l L_{n-l-1}^{2\,l+1}\left(\frac{2\,r}{n\,a_0}\right) \tag{17.17}$$

where we have defined the Bohr radius

$$a_0 = (r_0 \text{ for } E = E_1) = \sqrt{-\frac{\hbar^2}{2\,\mu\,E_1}} = \frac{\hbar^2}{\mu\,e^2} \tag{17.18}$$

and where we see that states of principal quantum number n have a scale radius $r_0 = n\,a_0$. $L_{n-l-1}^{2\,l+1}(2\,\rho)$ is the special polynomial of order $k = n - l - 1$ generated by the recursion relation, called the associated Laguerre polynomial. These polynomials have a similar recursion structure as the Legendre polynomials,

$$L_p^0 = e^x \frac{d^p}{dx^p}\left(e^{-x}x^p\right) \qquad L_p^k(x) = (-1)^k \frac{d^k}{dx^k} L_{p+k}^0 \tag{17.19}$$

The expectation value of the radial coordinate (the radius operator) is

$$\langle R \rangle_{n,l,m} = \frac{a_0}{2}\left[3\,n^2 - l(l+1)\right] \tag{17.20}$$

which can be found by doing the necessary integral directly or by a virial theorem argument given in Shankar. It is interesting to see that the expectation value of R depends on l even though the energy does not. There is a tradeoff between Coulomb energy and centrifugal barrier energy that allows $\langle R \rangle$ to change with l while leaving E unchanged.

"Accidental Degeneracy" in the Unperturbed Hydrogen Atom

We have the surprising lack of dependence of the energy eigenvalue E_n on the orbital quantum number l. We expect a lack of dependence on m from the spherical symmetry of the problem — neither the Coulomb potential nor the centrifugal barrier care how the wavefunction is oriented in angle, which is how m affects the state — but we do not obviously expect degeneracy in l since l appears in the radial equation! Why is there this degeneracy?

As we explained in Section 15.3, these situations arise via hidden symmetries, symmetries not obvious in the Hamiltonian. We may find generators for any such hidden symmetry and then define ladder operators that let us connect states that are degenerate in energy because of the hidden symmetry.

For the $1/r$ potential, classical mechanics tells us that, in addition to E and $\vec{l}$, there is an additional constant of the motion (conserved quantity) called the Laplace-Runge-Lenz vector,

$$\vec{n} = \frac{\vec{p} \times \vec{l}}{\mu} - \frac{e^2}{r} \vec{r} \qquad (17.21)$$

Recall that in the classical $1/r$ potential problem, the angular momentum is related the eccentricity of the orbit. There are multiple orbits in the same plane (and hence same angular momentum direction) that have the same total energy but different angular momentum magnitude and hence eccentricity. The classical symmetry transformation generated by the LRL vector implies that these orbits can be grouped into families that are related to each other by a transformation parameterized by a single parameter, which turns out to be the orbit eccentricity.

In detail: recall that the energy of an orbit is related to its eccentricity by

$$E = \frac{1}{2} \frac{k}{r_0} \left(\epsilon^2 - 1 \right) \tag{17.22}$$

($k = e^2$ in our case; see the Ph106ab Lecture Notes if you do not recall this result). For any energy, there is exactly one orbit that is perfectly circular ($\epsilon = 0$) with radius r_0, which is found by simply requiring circular motion where the centripetal force is provided by the Coulomb force. The above equation tells us that we can fix E while changing r_0 and ϵ in compensatory ways. Each member of the family of orbits of the same E may be labeled by its ϵ or its r_0. The LRL vector generates the coordinate transformation that transforms the circular orbit ($\epsilon = 0$) into any other orbit of the family ($\epsilon \neq 0$).

Let's construct the quantum mechanical operator analogue to $\vec{n}$. Symmetrizing the cross-product to obtain a Hermitian operator (as is needed to generate a unitary transformation), we have

$$\vec{N} = \frac{1}{2\,\mu} \left[\vec{P} \times \vec{L} - \vec{L} \times \vec{P} \right] - \frac{e^2}{\sqrt{X^2 + Y^2 + Z^2}} \, \vec{R} \tag{17.23}$$

The presence of position operators in the denominator is confusing and is really only valid in the position basis, where we interpret those operators as just being coordinate positions. The matrix elements of $\vec{N}$ in any other basis can be obtained by unitary transformation. We will not show explicitly that $[\vec{N}, H] = 0$, but one can show it most easily using the commutation relations (Section 13.4)

$$[R_i, P_j] = i\,\hbar\,\delta_{ij} \qquad\qquad [L_i, L_j] = \epsilon_{ijk}\,i\,\hbar\,L_k \tag{17.24}$$

$$[R_i, L_j] = \epsilon_{ijk}\,i\,\hbar\,R_k \qquad\qquad [P_i, L_j] = \epsilon_{ijk}\,i\,\hbar\,P_k \tag{17.25}$$

and also the usual rules about treating functions of operators via power series expansions; the commutators result in derivatives of these functions.

With the LRL Cartesian vector operator $\vec{N}$ in hand, we are of course inclined to write down the components of its spherical tensor representation, $\mathcal{N}^{(1)}$,

$$N_{2\pm1}^{(1)} = N_{\pm} = \mp \frac{1}{\sqrt{2}} \left[N_x \pm i\, N_y \right] \qquad N_2^{(1)} = N_0 = N_z \qquad (17.26)$$

Now, given a hydrogen atom state $|n, l, l\rangle$, what happens when we act on it with N_+? Recalling our discussion of spherical tensor operators in Section 15.3, it is the same as "adding angular momentum" with the state $|1, 1\rangle$:

$$N_+ |n, l, l\rangle \propto |1, 1\rangle \otimes |n, l, l\rangle = |n, l+1, l+1\rangle \qquad (17.27)$$

where we were able to make the last step because we know that, when adding angular momentum j_1 and j_2, the top state of the $j = j_1 + j_2$ ladder is obtained by the simple direct product of the highest states of the individual ladders, $|j_1, j_1\rangle \otimes |j_2, j_2\rangle$, because there are no other direct products with high enough m. This lets us connect states of the same n but different l. Then, states of the same l but different m are obtained by applying L_- as usual. Note that one can show that $N_+ |n, n-1, n-1\rangle = 0$ so that we do not violate the rules for which angular momentum states are allowed for a given n.

Other Comments

Please also review the remainder of Chapter 13 of Shankar — there is some good material on numerical estimates, relation to the Bohr model, and comparison to experiment. We will discuss multielectron atoms later.

Lecture 45:
The Perturbed Hydrogen Atom: The Stark Effect

Revision Date: 2008/02/29

Perturbed Hydrogen Atom: The Stark Effect

The Stark Effect

The first perturbation we consider for the hydrogen atom is the Stark Effect, which consists of applying a constant electric field in the $\hat{z}$ direction. We choose the $\hat{z}$ direction because that is the quantization axis for angular momentum and this fact will prove useful in doing the calculations of the perturbing matrix elements.

The classical Hamiltonian corresponding to the energy of the electron and proton in the applied electric field is

$$\delta \mathcal{H}^{(1)} = -e\,\phi(\vec{r}_e) + e\,\phi(\vec{r}_p) \tag{17.28}$$

$$= -e\,\vec{\mathcal{E}} \cdot [\vec{r}_p - \vec{r}_e] \qquad \text{(using } \vec{\mathcal{E}} = -\vec{\nabla}\phi) \tag{17.29}$$

$$= -e\,\vec{r} \cdot \vec{\mathcal{E}} \tag{17.30}$$

where $\vec{r} = \vec{r}_p - \vec{r}_e$ is the vector from the electron to the proton. This interaction is of the form

$$\delta \mathcal{H}^{(1)} = -\vec{d} \cdot \vec{\mathcal{E}} \qquad \text{with} \quad \vec{d} = e\,(\vec{r}_p - \vec{r}_e) \tag{17.31}$$

i.e., it is the electrical analog of the interaction of a magnetic moment with the magnetic field, with $\vec{d}$ being the electric dipole moment.

The resulting QM perturbing Hamiltonian is

$$\delta H^{(1)} = -e\,\mathcal{E}\,Z \tag{17.32}$$

where $\mathcal{E} = |\vec{\mathcal{E}}|$. We will have to later obtain a condition that establishes when the perturbation is small.

The first-order energy shift for any state is

$$E^{(1)}_{n,l,m} = e\,\mathcal{E}\langle n,l,m|Z|n,l,m\rangle \tag{17.33}$$

We may immediately see that this vanishes for *any* state, in two ways.

First we know that the Y_l^m are eigenstates of the parity operator with parity eigenvalue $(-1)^l$ (this is easy to show; the parity operation in spherical coordinates consists of the coordinate transformation $\theta \to \pi - \theta, \phi \to \phi + \pi$). The Z operator obviously has parity (-1), so the parity of $Z|n,l,m\rangle$ is $(-1)^{l+1}$. Explicitly:

$$\Pi\left(Z\,|n,l,m\rangle\right) = \Pi\,Z\,\Pi^{-1}\Pi|n,l,m\rangle = (-1)\,Z\,(-1)^{l+1}|n,l,m\rangle = (-1)^{l+1}Z\,|n,l,m\rangle \tag{17.34}$$

So the matrix element of $Z\,|n,l,m\rangle$ with $\langle n,l,m|$ vanishes because they are of opposite parity. This is the kind of parity selection rule we discussed in Section 16.3.

Perturbed Hydrogen Atom: The Stark Effect (cont.)

The second way to see this is via an angular momentum selection rule, or the Wigner-Eckart theorem. The Z operator is the $m_q = 0$ component of the spherical vector operator $\vec{R}$, $Z = R_2^{(1)}$. The projection theorem, which is a special case of the Wigner-Eckart theorem for vector operators, tells us

$$\langle n, l, m | Z | n, l, m \rangle = \langle n, l, m | R_2^{(1)} | n, l, m \rangle = \frac{\langle n, l, m | \vec{J} \cdot \vec{R} | n, l, m \rangle}{\hbar^2 j (j+1)} \langle l, m | (|1, 0 \rangle \otimes |l, m \rangle) \tag{17.35}$$

It is straightforward to see that the operator $\vec{J} \cdot \vec{R}$ vanishes:

$$\vec{J} \cdot \vec{R} = \left(\vec{R} \times \vec{P} \right) \cdot \vec{R} = \sum_{ijk} \epsilon_{ijk} R_j P_k R_i = \sum_{ijk} \epsilon_{ijk} R_i R_j P_k = 0 \tag{17.36}$$

We have used the usual ϵ_{ijk} notation to write out the cross-product. We then commute P_k and R_i because $\epsilon_{ijk} = 0$ when $i = k$; *i.e.*, we are assured $i \neq k$. Finally, ϵ_{ijk} is antisymmetric in i and k but $R_i R_k$ is symmetric in them, so $\sum_{ik} \epsilon_{ijk} R_i R_k = 0$.

Now we must calculate the second-order energy shift. Recall that our second formula for this shift, the one that does not require explicit calculation of the first-order eigenstate shift, is

$$E_{n,l,m}^{(2)} = e^2 \mathcal{E}^2 \sum_{n',l',m' \neq n,l,m} \frac{\left| \langle (n',l',m')^{(0)} | Z | (n,l,m)^{(0)} \rangle \right|^2}{E_n^{(0)} - E_{n'}^{(0)}} \qquad (17.37)$$

Note that no l index is required on the energies in the denominator due to the l degeneracy in the hydrogen atom.

The above formula has difficulties for $n > 1$ because there may be states for which $n = n'$ but $l \neq l'$ or $m \neq m'$ — *i.e.*, the contributing state is distinct from $|(n,l,m)^{(0)}\rangle$ but the principal quantum number is the same — in which case the denominator vanishes. This is the issue of doing perturbation theory when states are degenerate. So, the above formula is only valid for $n = 1$, for which there is no degeneracy. Let us consider only this case for now; we will study degenerate perturbation theory shortly.

Shankar shows how, for $n = 1$, one can bound the value of the above expression (his Method 1, p. 461). Note that his method would fail for $n > 1$, because it relies on the energy denominator always having the same sign, which is possible for $n = 1$ because $E_n^{(0)} < E_{n'}^{(0)}$ for all $n' \neq n$. But, since the formula is only valid for $n = 1$ anyways, it is a perfectly reasonable thing to do.

For the case $n = 1$, let's see if we can reduce the expression further. Let's evaluate the matrix element $\langle (n', l', m')^{(0)} | Z | (n, l, m)^{(0)} \rangle$. We may use the same kind of argument we made before using parity or angular momentum selection rules to see that the matrix element in general satisfies

$$\langle (n', l', m')^{(0)} | Z | (n, l, m)^{(0)} \rangle = \langle (n', l', m')^{(0)} | Z | (n, l, m)^{(0)} \rangle \, \delta_{l', l \pm 1} \delta_{m', m} \tag{17.38}$$

Therefore, our expression reduces to

$$E_{1,0,0}^{(2)} = e^2 \, \mathcal{E}^2 \sum_{n' \neq 1} \frac{\left| \langle (n', 1, 0)^{(0)} | Z | (1, 0, 0)^{(0)} \rangle \right|^2}{E_0^{(0)} - E_{n'}^{(0)}} \tag{17.39}$$

$$= \frac{e^2 \, \mathcal{E}^2}{[Ry]} \sum_{n' \neq 1} \frac{1 - (n')^2}{(n')^2} \left| \langle (n', 1, 0)^{(0)} | Z | (1, 0, 0)^{(0)} \rangle \right|^2 \tag{17.40}$$

Even with the restriction on the sum, this is not easy to evaluate because of the radial portion of the matrix element, which will be different for each n'.

Perturbed Hydrogen Atom: The Stark Effect (cont.)

Recall, however, the trick we developed in Section 16.1 for evaluating second-order energy shifts in general; if we can find an operator Ω such that

$$\delta H^{(1)} = \left[\Omega, H^{(0)}\right] \tag{17.41}$$

then the second-order energy shift reduces to

$$E_n^{(2)} = = \langle n^{(0)} | \delta H^{(1)} \Omega | n^{(0)} \rangle - \langle n^{(0)} | \delta H^{(1)} | n^{(0)} \rangle \langle n^{(0)} | \Omega | n^{(0)} \rangle \tag{17.42}$$

In our case, because the first-order shift vanishes and the second term is proportional to the first-order shift, the above would reduce to

$$E_n^{(2)} = \langle n^{(0)} | \delta H^{(1)} \Omega | n^{(0)} \rangle \tag{17.43}$$

Finding such an Ω is in general difficult, but we also saw that one can make the problem easier by only requiring

$$\delta H^{(1)} | n^{(0)} \rangle = \left[\Omega, H^{(0)}\right] | n^{(0)} \rangle \tag{17.44}$$

for the state $| n \rangle$ for which one wants to evaluate the energy shift. This specialization is fine here, since we know we can only evaluate the shift for $| 1, 0, 0 \rangle$ because of the degeneracies for higher states.

So we need to solve the equation

$$e\,\mathcal{E}\,Z|1,0,0\rangle = \left[\Omega, H^{(0)}\right]|1,0,0\rangle \tag{17.45}$$

Let's turn this into a differential equation for Ω in the position basis by applying $\langle\vec{r}|$ on the left:

$$\langle\vec{r}|e\,\mathcal{E}\,Z\,|1,0,0\rangle = \langle\vec{r}|\left[\Omega, H^{(0)}\right]|1,0,0\rangle \tag{17.46}$$

$$e\,\mathcal{E}\,z\,R_{10}(r)\sqrt{\frac{1}{4\pi}} = -\Omega_{\vec{r}}\,[Ry]\,R_{10}(r)\sqrt{\frac{1}{4\pi}} - H^{(0)}\,\Omega_{\vec{r}}\,R_{10}(r)\sqrt{\frac{1}{4\pi}} \tag{17.47}$$

$$e\,\mathcal{E}\,r\cos\theta\,R_{10}(r) = \left(\frac{\hbar^2}{2\,\mu}\nabla^2 + \frac{e^2}{r} - [Ry]\right)\Omega_{\vec{r}}\,R_{10}(r) \tag{17.48}$$

where $\Omega_{\vec{r}}$ is the position-basis representation of Ω and $1/\sqrt{4\pi}$ is Y_0^0. If one assumes Ω is a function of the coordinates only and not the momenta, then the above is a differential equation for $\Omega_{\vec{r}}(\vec{r})$.

The solution is

$$\Omega_{\vec{r}}(\vec{r}) = -\frac{\mu \, a_0 \, e \, \mathcal{E}}{\hbar^2} \left(\frac{r^2 \cos \theta}{2} + a_0 \, r \, \cos \theta \right) = -\frac{\mu \, a_0 \, e \, \mathcal{E}}{\hbar^2} \left(\frac{1}{2} \, |\vec{R}| + a_0 \right) Z \quad (17.49)$$

(This is not obvious!) With this operator, the second-order energy shift is

$$E_{1,0,0}^{(2)} = \langle (1,0,0)^{(0)} \, | \delta H^{(1)} \Omega | (1,0,0)^{(0)} \rangle \qquad (17.50)$$

$$= e \, \mathcal{E} \, \langle (1,0,0)^{(0)} \, | Z \, \Omega | (1,0,0)^{(0)} \rangle \qquad (17.51)$$

$$= -\frac{9}{4} \, a_0^3 \, \mathcal{E}^2 \qquad (17.52)$$

Again, this is not obvious, it requires some calculation, and the Wigner-Eckart theorem can be applied to do the angular integrals. Note that, because $\delta H^{(1)}$ and Ω each carry one power of Z, the operator whose matrix element must be taken is now parity $+1$ and so the matrix element does not vanish.

Now, that we have the second-order energy shift, we may check under what conditions perturbation theory is valid. The ratio of the second-order energy shift to the unperturbed energy is

$$\left| \frac{E_{1,0,0}^{(2)}}{E_{1,0,0}^{(0)}} \right| = \frac{\frac{9}{4} \, a_0^3 \, \mathcal{E}^2}{\frac{e^2}{2 \, a_0}} = \frac{9}{2} \, \frac{a_0^4 \, \mathcal{E}^2}{e^2} = \left(\frac{\sqrt{3}}{2} \, \frac{\mathcal{E}}{e/a_0^2} \right)^2 \qquad (17.53)$$

We have rewritten using the square because this is supposed to be a second-order energy correction. So, we need

$$\delta = \frac{\mathcal{E}}{e/a_0^2} \ll 1 \qquad (17.54)$$

in order for perturbation theory to be valid. This makes sense — the δ parameter is is just the ratio of the applied electric field to the electric field of the proton felt by the electron; the former must be weak in order for the hydrogen eigenstates to be "mostly" correct.

Now, let's try to physically interpret the energy shift. We expect it to be of the form $-\vec{d} \cdot \vec{\mathcal{E}}$ where $\vec{d}$ is the electric dipole moment. The dipole moment is proportional to the field, $\vec{d} \propto \vec{\mathcal{E}}$, because the field induces the dipole. This explains why the energy is quadratic in $\mathcal{E}$. What is the actual dipole moment? It is not just $|\vec{d}| = E_{1,0,0}^{(2)}/\mathcal{E}$ because the dipole moment changes with $\mathcal{E}$; rather, one must integrate the energy from $\mathcal{E} = 0$. Let's parameterize the dipole moment as $\vec{d} = \alpha \vec{\mathcal{E}}$. Then we have

$$U = -\int_0^{\mathcal{E}} \alpha \vec{\mathcal{E}}\,' \cdot d\vec{\mathcal{E}}\,' = -\int_0^{\mathcal{E}} \alpha \mathcal{E}\,'\, d\mathcal{E}\,' = -\frac{1}{2}\,\alpha\,\mathcal{E}^2 \tag{17.55}$$

So, the effective dipole moment is

$$\alpha\,\mathcal{E} = -2\,\frac{E_{1,0,0}^{(2)}}{\mathcal{E}} = -\frac{9}{2}\,a_0^3\,\mathcal{E} \tag{17.56}$$

The constant of proportionality is called the polarizability, and is

$$\alpha = -\frac{9}{2}\,a_0^3 \tag{17.57}$$

The energy shift is negative: the field creates a dipole moment that is aligned parallel to itself, which gives a negative energy according to the $-\vec{d} \cdot \vec{\mathcal{E}}$ form. The fact that the first-order energy shift vanishes can now be seen to be a result of the dipole moment being proportional to $\mathcal{E}$ because it is created by $\vec{\mathcal{E}}$ and the fact that the energy is linear in $\vec{d}$ and $\vec{\mathcal{E}}$.

Lecture 46:
Perturbed Hydrogen Atom: Fine Structure
Relativistic Kinetic Energy Correction

Revision Date: 2008/03/05

Perturbed Hydrogen Atom: Fine Structure

Hydrogen Fine Structure

We will apply time-independent perturbation theory, taking into account degeneracies, to calculate the first-order corrections due to two relativistic effects. The first effect includes relativistic corrections to the kinetic energy of the electron. The second effect is the magnetic dipole interaction of the electron's spin with the motional magnetic field of the proton seen by the electron due to the proton's motion in the electron's rest frame. This latter effect may not seem explicitly relativistic, but we will see that it depends on $(v/c)^2$ also, and so is indeed such an effect.

Relativistic Kinetic Energy Effect

We can determine the approximate size of v/c, which will determine how important relativistic corrections are, by calculating the ratio of the electron's KE to its rest mass. Let's just do a classical calculation for the hydrogen ground state to get a rough idea. The electron can be thought of as being in a circular orbit of radius a_0, and so its KE is $e^2/2\,a_0$. Therefore,

$$\left(\frac{v}{c}\right)^2 = \frac{\mu\,v^2}{\mu\,c^2} = 2\,\frac{KE}{\mu\,c^2} = \frac{e^2/a_0}{\mu\,c^2} = \frac{2 \times 13.6 \text{ eV}}{511 \text{ keV}} \approx 5 \times 10^{-5} \tag{17.58}$$

The term of order $(v/c)^2$ is already being included. But the above implies that the next-order term, of order $(v/c)^4$, will provide a correction of the above size to the nonrelativistic energy.

Let's calculate the kinetic energy to order $(v/c)^4$:

$$T = (\gamma - 1)\,\mu\,c^2 = \left(\left(1 - (v/c)^2\right)^{-1/2} - 1 \right) \mu\,c^2 \tag{17.59}$$

$$= \left[\frac{1}{2}\left(\frac{v}{c}\right)^2 - \frac{1}{8}\left(\frac{v}{c}\right)^4 \right] \mu\,c^2 \tag{17.60}$$

$$= \frac{P^2}{2\,\mu} - \frac{P^4}{8\,\mu^3\,c^2} \tag{17.61}$$

Now, let's apply (degenerate) perturbation theory to this additional term,
$H_T = -P^4/8\,\mu^3\,c^2$.

The first thing we notice is that P^4 is a scalar operator. This is great, as it means that, in the degenerate subspace $\mathcal{D}_n$, it is already diagonalized: because it is a scalar, it adds no angular momentum, so it can only have nonzero matrix elements among states of the same l, m:

$$\langle n', l', m' | P^4 | n, l, m \rangle = \langle n', l, m | P^4 | n, l, m \rangle\, \delta_{l'l}\delta_{m'm} \tag{17.62}$$

So, essentially, we have already done the diagonalization step necessary for degenerate perturbation theory. Now we proceed as usual.

The first-order energy shift is

$$E_{T,n}^{(1)} = -\frac{1}{8\,\mu^3\,c^2}\,\langle n, l, m | P^4 | n, l, m \rangle \tag{17.63}$$

We evaluate the matrix element by using the fact

$$P^4 = 4\,\mu^2\left(\frac{P^2}{2\,\mu}\right) = 4\,\mu^2\left(H^{(0)} + \frac{e^2}{r}\right)^2 \tag{17.64}$$

So we have

$$E_{T,n}^{(1)} = -\frac{1}{2\,\mu\,c^2}\,\langle n, l, m |\left(H^{(0)} + \frac{e^2}{r}\right)^2 | n, l, m \rangle \tag{17.65}$$

$$= -\frac{1}{2\,\mu\,c^2}\left(\left[E_n^{(0)}\right]^2 + 2\,E_n^{(0)}\,e^2\left\langle \frac{1}{r}\right\rangle_{n,l,m} + e^4\left\langle \frac{1}{r^2}\right\rangle_{n,l,m}\right) \tag{17.66}$$

Perturbed Hydrogen Atom: Fine Structure (cont.)

The expectation values can be calculated easily using two clever tricks:

▶ The first term is just $-\langle V \rangle$. Presumably, you have all seen the virial theorem in classical mechanics. (If not, see Section 1.3.2 of the Ph106ab Lecture Notes; the derivation relies only on Newtonian mechanics for systems of particles.) For $1/r$ potentials, the virial theorem is

$$\langle T \rangle = -\frac{1}{2} \langle V \rangle \tag{17.67}$$

where the angle brackets indicate time averaging. Ehrenfest's theorem tells us the virial theorem will be true for expectation values of QM operators (this can be proven rigorously, see Shankar 13.1.5). Therefore, we know

$$e^2 \left\langle \frac{1}{r} \right\rangle_{n,l,m} = -\langle V \rangle_{n,l,m} = -2 \left(\langle T \rangle_{n,l,m} + \langle V \rangle_{n,l,m} \right) \tag{17.68}$$

$$= -2 \left\langle H^{(0)} \right\rangle_{n,l,m} = -2\, E_n^{(0)} \tag{17.69}$$

Perturbed Hydrogen Atom: Fine Structure (cont.)

▶ The second term is of the same form as the centrifugal term in the radial equation, so it changes that term from $l(l + 1)\,\hbar^2/2\,\mu\,r^2$ to $\left[l(l+1)\hbar^2/2\,\mu + e^4\right]/r^2$. If we define l' by

$$\frac{l\,(l+1)\,\hbar^2}{2\,\mu\,r^2} + \frac{\lambda}{r^2} = \frac{l'\,(l'+1)\,\hbar^2}{2\,\mu\,r^2} \tag{17.70}$$

($\lambda = e^4$ in our case) then we just repeat the derivation of the unperturbed hydrogen atom with l replaced by $l'(\lambda)$. The energy becomes

$$E(\lambda) = E_1^{(0)} \frac{1}{(n + (l'(\lambda) - l))^2} \tag{17.71}$$

We Taylor expand around $\lambda = 0$ to obtain

$$\left\langle \frac{\lambda}{r^2} \right\rangle_{n,l,m} = E(\lambda) - E(\lambda = 0) \tag{17.72}$$

$$= E_1^{(0)} \left[-\frac{2}{n^3} \right] \frac{dl'}{d\lambda}\,\lambda = -\frac{2\,E_n^{(0)}}{n} \frac{2\,\mu}{\hbar^2\,(2\,l+1)}\,\lambda \tag{17.73}$$

So, we have

$$\left\langle \frac{e^4}{r^2} \right\rangle_{n,l,m} = \frac{4\left[E_n^{(0)}\right]^2 n}{l + 1/2} \tag{17.74}$$

Perturbed Hydrogen Atom: Fine Structure (cont.)

With these tricks, our result is

$$E_{T,n}^{(1)} = -\frac{1}{2\,\mu\,c^2} \left(\left[E_n^{(0)}\right]^2 - 4\left[E_n^{(0)}\right]^2 + \frac{4\left[E_n^{(0)}\right]^2 n}{l + 1/2} \right) \tag{17.75}$$

$$= -\frac{[E_n^{(0)}]^2}{2\,\mu\,c^2} \left(-3 + \frac{4\,n}{l + 1/2} \right) \tag{17.76}$$

We may write this in a different, illustrative form. $\alpha = e^2/\hbar\,c \approx 1/137$ is the fine structure constant and characterizes the strength of these relativistic corrections to the hydrogen atom energies. In terms of α, the hydrogen atom unperturbed energies are

$$E_n^{(0)} = -\frac{1}{2}\,\frac{(\mu\,c^2)\,\alpha^2}{n^2} \tag{17.77}$$

so the correction we have just derived may be written

$$E_{T,n}^{(1)} = -\frac{1}{2}\,(\mu\,c^2)\,\alpha^4 \left(-\frac{3}{4\,n^4} + \frac{1}{n^3\,(l + 1/2)} \right) \tag{17.78}$$

Lecture 47:
Perturbed Hydrogen Atom: Fine Structure
Spin-Orbit Correction
Total Correction
Spectroscopic Notation

Revision Date: 2008/03/05

Perturbed Hydrogen Atom: Fine Structure

Spin-Orbit Interaction

Let's calculate this effect in the rest frame of the electron. If the electron is moving at velocity $\vec{v}$ in the lab frame, then the proton moves at velocity $-\vec{v}$ in the electron's rest frame and the magnetic field due to the moving proton is

$$\vec{B} = -\frac{e}{c} \frac{\vec{v} \times \vec{r}}{r^3} \tag{17.79}$$

Another way to obtain this expression is to recall that, according to the Lorentz transformation properties of the EM field, an electric field $\vec{\mathcal{E}}$ in one reference frame results in a magnetic field $-\frac{\vec{v}}{c} \times \vec{\mathcal{E}}$ in a frame moving at velocity $\vec{v}$ with respect to the first (see, *e.g.*, *Classical Electrodynamics* by Jackson). This second form is more useful because one can use it for many-electron atoms in which the electric field seen by the electron is not a simple Coulomb field; in that case, the magnetic field seen by the electron is

$$\vec{B} = \frac{1}{c} \frac{\vec{v} \times \vec{r}}{r} \frac{dV(r)}{dr} \tag{17.80}$$

where $V(r)$ is the electrostatic potential in the lab frame experience by the electron. We shall restrict to $V(r) = -e^2/r$ for simplicity, here, though.

In the electron's rest frame, it has no orbital angular momentum, so we need only consider the interaction of its spin angular momentum with the magnetic field, which is, as usual,

$$H_{so} = - \left(I^{(o)} \otimes \vec{\mu}^{(s)} \right) \cdot \left(\vec{B}^{(o)} \otimes I^{(s)} \right) \tag{17.81}$$

$$= - \left(-\frac{e}{\mu c} I^{(o)} \otimes \vec{S}^{(s)} \right) \cdot \left(\frac{e}{c} \frac{1}{r^3} \frac{\vec{L}^{(o)} \otimes I^{(s)}}{\mu} \right) \tag{17.82}$$

$$= \frac{e^2}{\mu^2 c^2 r^3} \left[I^{(o)} \otimes \vec{S}^{(s)} \right] \cdot \left[\vec{L}^{(o)} \otimes I^{(s)} \right] \tag{17.83}$$

where the $^{(o)}$ and $^{(s)}$ indicate which Hilbert space each operator lives in ("orbital" for the spatial state and "spin" for the spin state).

We are of course considering the direct product $\mathcal{V} \otimes V^{(1/2)}$ that takes into account both the orbital and spin degrees of freedom in the hydrogen atom. Note that Shankar uses the shorthand $\vec{S} \cdot \vec{L}$ for $\left[I^{(o)} \otimes \vec{S}^{(s)} \right] \cdot \left[\vec{L}^{(o)} \otimes I^{(s)} \right]$. We have also used the fact that the electron's gyromagnetic ratio is 2 to a very good approximation.

When one transforms back to the lab frame, it can be shown that one acquires a factor of 1/2 out front due to the fact that the electron is not moving at fixed velocity. A more rigorous way to obtain this factor of 1/2 is to use the fully relativistic Dirac model for the electron so that one need not go into the electron's rest frame to obtain the magnetic part of its interaction with the proton. (The Coulomb interaction is the nonrelativistic limit of the interaction of the electron with a static EM field.)

Now, we proceed as usual with degenerate perturbation theory. First, we must diagonalize H_{so} in each $\mathcal{D}_n$ subspace. Note that, now that $\vec{S}$ is included in the Hamiltonian, we must also consider the spin state, label the states by $|n, l, m, m_s \rangle$, and note this additional two-fold degeneracy in $H^{(0)}$ at each n, l, m value.

We could work directly with $\left[I^{(o)} \otimes \vec{S}^{(s)} \right] \cdot \left[\vec{L}^{(o)} \otimes I^{(s)} \right]$, but we will find things are much simpler if we instead rewrite H_{so} as follows (including the correct prefactor of 1/2 as explained above):

$$H_{so} = \frac{1}{2} \frac{e^2}{\mu^2 c^2 r^3} \left[I^{(o)} \otimes \vec{S}^{(s)} \right] \cdot \left[\vec{L}^{(o)} \otimes I^{(s)} \right] \tag{17.84}$$

$$= \frac{e^2}{4 \mu^2 c^2 r^3} \left[J^2 - (L^2)^{(o)} \otimes I^{(s)} - I^{(o)} \otimes (S^2)^{(s)} \right] \tag{17.85}$$

which follows from $\vec{J} = \vec{L}^{(o)} \otimes I^{(s)} + I^{(o)} \otimes \vec{S}^{(s)}$.

Now, the degenerate subspace $\mathcal{D}_n$ consists of the states $|\, n, l, m, m_s \,\rangle$ for the particular value of n and for all allowed values of l, m, and m_s. To be explicit, we are considering the space (uncoupled representation)

$$V = \left(V^{(0)} \oplus V^{(1)} \oplus \cdots \oplus V^{(n-2)} \oplus V^{(n-1)} \right) \otimes V^{(1/2)} \tag{17.86}$$

which decomposes in terms of direct sums of total angular momenta as (coupled representation)

$$V = V^{(1/2)}_{(0)} \oplus V^{(1/2)}_{(1)} \oplus V^{(3/2)}_{(1)} \oplus V^{(3/2)}_{(2)} \oplus \cdots \oplus V^{(n-3/2)}_{(n-2)} \oplus V^{(n-3/2)}_{(n-1)} \oplus V^{(n-1/2)}_{(n-1)} \tag{17.87}$$

where we have labeled the spaces with a superscript that, as usual, says what the total angular momentum is, and in addition a subscript that tells us which l the given subspace came from. We obtain two copies of each subspace $V^{(j)}$ because we get one from $V^{(j-1/2)} \otimes V^{(1/2)}$ and one from $V^{(j+1/2)} \otimes V^{(1/2)}$, excluding the special case of $j = n - 1/2$.

We now can see that, when we work in the coupled representation, H_{so} will be diagonalized in the degenerate subspace:

$$\langle n, j', m'; l', s = 1/2 \, |H_{so}| n, j, m_j; l, s = 1/2 \, \rangle$$

$$= \delta_{j'j} \delta_{m'm} \delta_{l'l} \frac{e^2}{4\,\mu^2\,c^2} \left\langle \frac{1}{r^3} \right\rangle_{n,l} \hbar^2 \left[j\,(j+1) - l\,(l+1) - 3/4 \right] \qquad (17.88)$$

Note that m_j gives the eigenvalue of J_z, not of L_z. We have done what we needed to do, which is to find a basis for each degenerate subspace in which the perturbing Hamiltonian is diagonal. The above matrix elements are therefore the first-order energy shifts as a function of j and l. Note that n is the same for both states because we only need to diagonalize H_{so} in a given degenerate subspace; we don't need to worry yet about the matrix elements of H_{so} between states that are nondegenerate; *i.e.*, have different n.

To be explicit, the energy shifts are

$$V^{(j)}_{l=j-1/2}: \quad E^{(1)}_{so} = \frac{\hbar^2 e^2}{4\,\mu^2 c^2} \left\langle \frac{1}{r^3} \right\rangle_{n,l} l \tag{17.89}$$

$$V^{(j)}_{l=j+1/2}: \quad E^{(1)}_{so} = \frac{\hbar^2 e^2}{4\,\mu^2 c^2} \left\langle \frac{1}{r^3} \right\rangle_{n,l} (-1)(l+1) \tag{17.90}$$

where we show the energy shifts for states in the subspaces $V^{(j)}_{l=j-1/2}$ and $V^{(j)}_{l=j+1/2}$.

Perturbed Hydrogen Atom: Fine Structure (cont.)

Next, we need to evaluate $\left\langle \frac{1}{r^3} \right\rangle_{n,l}$. This requires another trick (Shankar 17.3.4), wherein one defines the radial momentum operator

$$p_r = \frac{i}{\hbar} \left(\frac{\partial}{\partial r} + \frac{1}{r} \right) \tag{17.91}$$

calculates $[H^{(0)}, p_r]$ explicitly, and makes use of $\langle [H^{(0)}, p_r] \rangle = 0$ for energy eigenstates (but note that $[H^{(0)}, p_r] \neq 0$!) to obtain the relation

$$\left\langle \frac{1}{r^3} \right\rangle_{n,l} = \frac{1}{a_0 l (l+1)} \left\langle \frac{1}{r^2} \right\rangle_{n,l} \tag{17.92}$$

and where we already have a means to evaluate $\left\langle \frac{1}{r^2} \right\rangle_{n,l}$. The result is

$$\left\langle \frac{1}{r^3} \right\rangle_{n,l} = \frac{1}{a_0^3} \frac{1}{n^3} \frac{1}{l (l+1/2) (l+1)} \tag{17.93}$$

Inserting this into our expression for the energy shifts, we obtain

$$V_{l=j-1/2}^{(j)}: \quad E_{so}^{(1)} = \frac{1}{4} \left(\mu c^2 \right) \alpha^4 \frac{1}{n^3} \frac{1}{(l+1/2)(l+1)} \tag{17.94}$$

$$= \frac{1}{4} \left(\mu c^2 \right) \alpha^4 \frac{1}{n^3} \frac{1}{(l+1/2)(j+1/2)} \tag{17.95}$$

$$V_{l=j+1/2}^{(j)}: \quad E_{so}^{(1)} = \frac{1}{4} \left(\mu c^2 \right) \alpha^4 \frac{1}{n^3} \frac{-1}{l(l+1/2)} \tag{17.96}$$

$$= \frac{1}{4} \left(\mu c^2 \right) \alpha^4 \frac{1}{n^3} \frac{-1}{(l+1/2)(j+1/2)} \tag{17.97}$$

This is our result for the spin-orbit energy shift. We see that the degeneracy in l and j is broken, but there is still degeneracy in m (the J_z eigenvalue). Instead of n^2 degenerate states at each n, there are $2j+1$ degenerate states at each $(j, l = j \pm 1/2)$ pair. Note that states of the same l can yield two different j values and the same j can arise from two different l values, but that there is no degeneracy there: each $j, l = j \pm 1/2$ pair corresponds to a unique energy shift. Another way of saying this is that each $V_{l=j\pm1/2}^{(j)}$ subspace is degenerate with no other, but of course all the states inside the $2j+1$-dimensional $V_{l=j\pm1/2}^{(j)}$ subspace are degenerate (the m degeneracy).

Total Fine Structure Correction

Adding together with the relativistic kinetic energy shift, we get the total fine structure energy shift:

$$E_{fs,(n,j,l)}^{(1)} = E_{T,(n,l)}^{(1)} + E_{so,(n,j,l=j\pm 1/2)}^{(1)} \tag{17.98}$$

$$= -\frac{1}{2} \frac{\mu c^2 \alpha^2}{n^2} \frac{\alpha^2}{n} \left(-\frac{3}{4n} + \frac{1}{l+1/2} \mp \frac{1}{2} \frac{1}{(l+1/2)(j+1/2)} \right) \tag{17.99}$$

$$= -\frac{1}{2} \frac{\mu c^2 \alpha^2}{n^2} \frac{\alpha^2}{n} \left(-\frac{3}{4n} + \frac{j+1/2 \mp 1/2}{(l+1/2)(j+1/2)} \right) \tag{17.100}$$

$$= -\frac{1}{2} \frac{\mu c^2 \alpha^2}{n^2} \frac{\alpha^2}{n} \left(-\frac{3}{4n} + \frac{j+1/2 \mp 1/2}{(j \mp 1/2 + 1/2)(j+1/2)} \right) \tag{17.101}$$

$$= -\frac{1}{2} \frac{\mu c^2 \alpha^2}{n^2} \frac{\alpha^2}{n} \left(\frac{1}{j+1/2} - \frac{3}{4n} \right) \tag{17.102}$$

where we used $l = j \mp 1/2$ in the above. The relativistic nature of this correction is evidence by the fact that it depends on μc^2 and α in the same way as the relativistic KE correction.

Recall that $E_n^{(0)} = -\frac{1}{2}(\mu c^2) \alpha^2 / n^2$, so we may also write the above as

$$\frac{E_{fs,(n,j,l)}^{(1)}}{|E_n^0|} = -\frac{\alpha^2}{n}\left(\frac{1}{j+1/2} - \frac{3}{4\,n}\right) \tag{17.103}$$

The fact that the two corrections' l-dependence is of the same form and results in a simplified j-dependent term with no explicit l dependence is interesting. First, it immediately tells us that the two $V_{l=j\pm 1/2}^{(j)}$ with the same j but different l are degenerate, unlike when we considered the spin-orbit interaction alone. Therefore, there is a $2 \times (2j+1)$ degeneracy at each j (except the highest j for a given n, $j = n - 1/2$, which is only $2j + 1$-degenerate). The simplified form one obtains when one combines the two results also corroborates the idea that the two effects are both relativistic effects. When we consider the fully relativistic Dirac theory of the electron, one expects the two to fall out immediately and together.

We plot the fine structure corrections for $n = 1$ and $n = 2$ on the following page. The complete breaking of n, l degeneracy by the spin-orbit interaction, followed by the partial recovery of degeneracy when the KE and spin-orbit corrections are summed, is visible, both in the final result and the relative size and signs of the KE and spin-orbit corrections for different l at the same j.
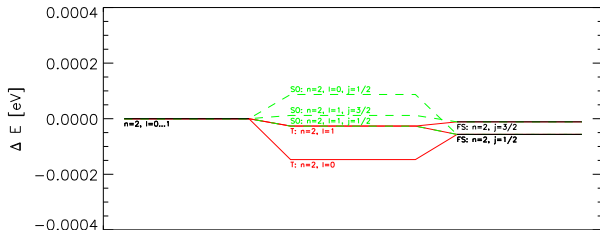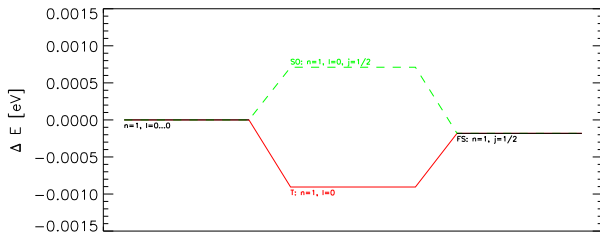
Perturbed Hydrogen Atom: Fine Structure (cont.)

Spectroscopic Notation

We may now introduce the Russell-Saunders term notation,

$$^{2s+1}L_j \qquad (17.104)$$

The L symbol will be $S, P, D, F, G, \ldots$ for $l = 0, 1, 2, 3, 4, \ldots$. The bottom symbol gives the total angular momentum quantum number. The superscript is a bit more complicated. It is intended to indicate the level of degeneracy of each j level after the fine-structure corrections are made. This is set by the number of different l levels that can yield a particular value of j. This is set by the electron spin; there are $2s + 1 = 2$ different l levels that can yield a given j. So the superscript is always 2 for single-electron atoms. The superscript becomes more useful when considering multielectron atoms, when the total spin can be larger than $s = 1/2$. For single-electron atoms, the above will always be of the form

$$^{2}L_{l\pm1/2} \qquad (17.105)$$

So, for the two diagrams shown on the previous page, we have:

$$n = 1, l = 0, j = 1/2 : \ ^2S_{1/2} \tag{17.106}$$

$$n = 2, l = 1, j = 3/2 : \ ^2P_{3/2} \tag{17.107}$$

$$n = 2, l = 1, j = 1/2 : \ ^2P_{1/2} \tag{17.108}$$

$$n = 2, l = 0, j = 1/2 : \ ^2S_{1/2} \tag{17.109}$$

It is clear that spectroscopic notation is somewhat ambiguous. It does not distinguish the $n = 1$ and $n = 2$ versions of $^2S_{1/2}$; it is assumed that which principal quantum number is relevant will be clear from context. Note that the $^2P_{1/2}$ and $^2S_{1/2}$ levels for $n = 2$ are degenerate in energy even after fine-structure corrections are included because their j values are the same.

Section 18
Multielectron Atoms

Lecture 48:
The Helium Atom

Revision Date: 2008/03/10

The Helium Atom

Overview

Two-electron atoms are the clear next case to consider. There are two new effects that appear:

▶ **Exchange Effects**
  Because the two electrons in the atom are identical fermionic particles, one must require the Hilbert space state for the atom be antisymmetric under exchange of the two electrons. This has interesting consequences because it ties together the spin and spatial parts of the state and results in an effective repulsion between the electrons when they are in the same spin state.

▶ **Electrostatic Repulsion**
  We must account for the fact that the two electrons in the atom electrostatically repel each other and that this affects the energy.

The Helium Atom (cont.)

Two-Electron States and Antisymmetry under Exchange

In the absence of electrostatic repulsion or antisymmetrization issues, the Hilbert space available to a two-electron atoms is simply the direct product of the hydrogen atom Hilbert space with itself:

$$|\psi\rangle = |n_1, l_1, m_1, m_{s,1}\rangle \otimes |n_2, l_2, m_2, m_{s,2}\rangle \tag{18.1}$$

The Hamiltonian for a two-electron atom is a separable sum of the Hamiltonians for each electron in the Coulomb field of the nucleus:

$$H_{2e} = H_{1e}^{(1)} \otimes I^{(2)} + I^{(1)} \otimes H_{1e}^{(2)} \tag{18.2}$$

where H_{1e} is just a one-electron Hamiltonian and, as usual, the $^{(n)}$ superscript indicates which factor space the operator works in.

However, because the electrons are identical fermions and have access to the same states in their own factor Hilbert spaces, we must restrict ourselves to states that are antisymmetric under exchange of the two electrons. So let's build a basis of states that satisfy this condition.

The Helium Atom (cont.)

Let's begin by considering the allowed spin states (the spin factor space of each particle's Hilbert space). Before symmetrization, they are

$$\left| \frac{1}{2}, \frac{1}{2} \right\rangle^{(1)} \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle^{(2)} \qquad \left| \frac{1}{2}, \frac{1}{2} \right\rangle^{(1)} \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle^{(2)} \tag{18.3}$$

$$\left| \frac{1}{2}, -\frac{1}{2} \right\rangle^{(1)} \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle^{(2)} \qquad \left| \frac{1}{2}, -\frac{1}{2} \right\rangle^{(1)} \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle^{(2)} \tag{18.4}$$

Note that we don't necessarily care if we are building eigenstates of the total spin magnitude and z-projection operators, so it is ok to stay with the direct product basis.

The Helium Atom (cont.)

Now let's find the symmetrized and antisymmetrized combinations. *Note that these states will no longer be eigenstates of the unsymmetrized spin operators,* $(S^2)^{(1)} \otimes I^{(2)}, S_z^{(1)} \otimes I^{(2)}, I^{(1)} \otimes (S^2)^{(2)}, I^{(1)} \otimes S_z^{(1)}$. *Recall that we had to construct symmetrized operators when dealing with identical particles (Section 8.4).* The symmetrized states are

$$\left| \frac{1}{2}, \pm\frac{1}{2}; \frac{1}{2}, \pm\frac{1}{2} \right\rangle_+^{(1)\otimes(2)} = \left| \frac{1}{2}, \pm\frac{1}{2} \right\rangle^{(1)} \otimes \left| \frac{1}{2}, \pm\frac{1}{2} \right\rangle^{(2)} \tag{18.5}$$

$$\left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle_+^{(1)\otimes(2)} = \frac{1}{\sqrt{2}} \left(\left| \frac{1}{2}, \frac{1}{2} \right\rangle^{(1)} \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle^{(2)} + \left| \frac{1}{2}, -\frac{1}{2} \right\rangle^{(1)} \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle^{(2)} \right) \tag{18.6}$$

$$\left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle_-^{(1)\otimes(2)} = \frac{1}{\sqrt{2}} \left(\left| \frac{1}{2}, \frac{1}{2} \right\rangle^{(1)} \otimes \left| \frac{1}{2}, -\frac{1}{2} \right\rangle^{(2)} - \left| \frac{1}{2}, -\frac{1}{2} \right\rangle^{(1)} \otimes \left| \frac{1}{2}, \frac{1}{2} \right\rangle^{(2)} \right) \tag{18.7}$$

Note that there are no $\left| \frac{1}{2}, \pm\frac{1}{2}; \frac{1}{2}, \pm\frac{1}{2} \right\rangle_-^{(1)\otimes(2)}$ states — they would vanish, and they would also violate the dimensionality of the space (without them, we already have enough basis elements).

Now, it is clear that the states we have constructed also happen to be the ones we construct when we "add angular momentum" by constructing a basis for the direct product space consisting of eigenstates of the total spin angular momentum squared and z-projection operators,

$$(S^2)^{(1)\otimes(2)} = \left(\vec{S}^{(1)} \otimes I^{(2)} + I^{(1)} \otimes \vec{S}^{(2)} \right)^2 \tag{18.8}$$

$$= (S^2)^{(1)} \otimes I^{(2)} + I^{(1)} \otimes (S^2)^{(2)} + 2\, \vec{S}^{(1)} \dot{\otimes} \vec{S}^{(2)} \tag{18.9}$$

$$S_z^{(1)} \otimes (2) = S_z^{(1)} \otimes I^{(2)} + I^{(1)} \otimes S_z^{(2)} \tag{18.10}$$

This conveniently occurs because these operators are symmetric under exchange of the two particles *and* because both particles have $s = 1/2$. (We shall explain this further below.)

So we have

$$| S = 1, M_S = \pm 1 \rangle = \left| \frac{1}{2}, \pm \frac{1}{2}; \frac{1}{2}, \pm \frac{1}{2} \right\rangle_{+}^{(1)\otimes(2)} \tag{18.11}$$

$$| S = 1, M_S = 0 \rangle = \left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle_{+}^{(1)\otimes(2)} \tag{18.12}$$

$$| S = 0, M_S = 0 \rangle = \left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle_{-}^{(1)\otimes(2)} \tag{18.13}$$

where S and M_S label the eigenvalues of $(S^2)^{(1)\otimes(2)}$ and $S_z^{(1)\otimes(2)}$.

When we combine the spin and spatial pieces of the state, we will need the spatial piece to carry the opposite symmetry as the spin piece to ensure the direct product state is antisymmetric under exchange. So we need to construct symmetrized and antisymmetrized basis elements for the spatial piece of the two-particle system. Let's try to do this in the same way as we initially built the basis elements for the spin space:

$$
\begin{aligned}
&|n_1, l_1, m_1; n_2, l_2, m_2 \rangle_{\pm}^{(1)\otimes(2)} \\
&= \frac{1}{\sqrt{2}} \left(|n_1, l_1, m_1 \rangle^{(1)} \otimes |n_2, l_2, m_2 \rangle^{(2)} \pm |n_2, l_2, m_2 \rangle^{(1)} \otimes |n_1, l_1, m_1 \rangle^{(2)} \right)
\end{aligned} \quad (18.14)
$$

While these symmetrized/antisymmetrized states are a reasonable basis, they are not eigenstates of total orbital angular momentum like we found in the spin case; it would be convenient if they were.

The Helium Atom (cont.)

To see this last point clearly, recall that the total orbital angular momentum operators are

$$(L^2)^{(1)\otimes(2)} = \left(\vec{L}^{(1)} \otimes I^{(2)} + I^{(1)} \otimes \vec{L}^{(2)}\right)^2 \tag{18.15}$$

$$= (L^2)^{(1)} \otimes I^{(2)} + I^{(1)} \otimes (L^2)^{(2)} + 2\,\vec{L}^{(1)}\dot{\otimes}\vec{L}^{(2)} \tag{18.16}$$

$$L_z^{(1)\otimes(2)} = L_z^{(1)} \otimes I^{(2)} + I^{(1)} \otimes L_z^{(2)} \tag{18.17}$$

because the $\vec{L}^{(1)}\dot{\otimes}\vec{L}^{(2)}$ term contains $L_{\pm}^{(1)}L_{\mp}^{(2)}$, which yield states of higher and lower m_1 and m_2 that cause problems unless $l_1 = l_2$ and $|m_1 + m_2| \geq 2\,l_1 - 1$. This condition can only be met for all allowed m_1, m_2 by $l_1 = l_2 = 1/2$, which is not allowed for orbital angular momentum.

The Helium Atom (cont.)

It would be awfully convenient if we could construct symmetrized/antisymmetrized states that are eigenstates of total orbital angular momentum, so let's try another technique: let's start with eigenstates of the total orbital angular momentum operators and then symmetrize them. By default, eigenstates of total orbital angular momentum are not symmetric or antisymmetric under interchange because $l_1 \neq l_2$ in general. But it's easy to construct symmetrized and antisymmetrized states. Suppose we have a state

$$
|n_1, n_2, l, m, l_1, l_2\rangle = \sum_{m_1 = -l_1}^{l_1} \left(\langle n_1, l_1, m_1 |^{(1)} \otimes \langle n_2, l_2, m - m_1 |^{(2)} \right) |n_1, n_2, l, m, l_1, l_2\rangle \\
\times |n_1, l_1, m_1\rangle^{(1)} \otimes |n_2, l_2, m - m_1\rangle^{(2)} \qquad (18.18)
$$

where we take the direct product state to vanish if $|m - m_1| > l_2$.

Then we can easily construct the symmetrized and antisymmetrized states

$$
\begin{aligned}
&|(n_1, n_2), l, m, (l_1, l_2)\rangle_\pm \\
&= \sum_{m_1 = -l_1}^{l_1} \left(\langle n_1, l_1, m_1 |^{(1)} \otimes \langle n_2, l_2, m - m_1 |^{(2)} \right) |n_1, n_2, l, m, l_1, l_2\rangle \\
&\qquad \frac{1}{\sqrt{2}} \left(|n_1, l_1, m_1\rangle^{(1)} \otimes |n_2, l_2, m - m_1\rangle^{(2)} \pm |n_2, l_2, m - m_1\rangle^{(1)} \otimes |n_1, l_1, m_1\rangle^{(2)} \right)
\end{aligned}
\tag{18.19}
$$

We put parentheses around n_1, n_2 and l_1, l_2 to indicate that we have symmetrized or antisymmetrized in these quantum numbers and that therefore the state is no longer an eigenstate of the single-particle operators corresponding to these quantum numbers $(H_{1e}^{(1)} \otimes I^{(2)}, I^{(1)} \otimes H_{1e}^{(2)}, (L^2)^{(1)} \otimes I^{(2)}, I^{(1)} \otimes (L^2)^{(2)})$. These states *are* eigenstates of $H_{2e}, (L^2)^{(1) \otimes (2)}$, and $L_z^{(1) \otimes (2)}$.

The Helium Atom (cont.)

The above expression is incorrect when $n_1 = n_2$, $l_1 = l_2$. In these cases, one should just use the unsymmetrized version of the expression, so that the full expression should be

$$
\begin{aligned}
&|(n_1, n_2), l, m, (l_1, l_2)\,\rangle_{\pm} \\
&= \sum_{m_1 = -l_1}^{l_1} \left(1 - \delta_{n_1 n_2}\delta_{l_1 l_2}\right) \left(\langle n_1, l_1, m_1\,|^{(1)} \otimes \langle n_2, l_2, m - m_1\,|^{(2)}\right) |n_1, n_2, l, m, l_1, l_2\,\rangle \\
&\qquad\qquad \frac{1}{\sqrt{2}} \left(|n_1, l_1, m_1\,\rangle^{(1)} \otimes |n_2, l_2, m - m_1\,\rangle^{(2)} \pm |n_2, l_2, m - m_1\,\rangle^{(1)} \otimes |n_1, l_1, m_1\,\rangle^{(2)}\right) \\
&\quad + \sum_{m_1 = -l_1}^{l_1} \delta_{n_1 n_2}\delta_{l_1 l_2} \left(\langle n_1, l_1, m_1\,|^{(1)} \otimes \langle n_2, l_2, m - m_1\,|^{(2)}\right) |n_1, n_2, l, m, l_1, l_2\,\rangle \\
&\qquad\qquad |n_1, l_1, m_1\,\rangle^{(1)} \otimes |n_2, l_2, m - m_1\,\rangle^{(2)}
\end{aligned}
\tag{18.20}
$$

The form is unchanged unless $n_1 = n_2$, $l_1 = l_2$. When these two conditions are met, we get to use the unsymmetrized expression because the states are already symmetrized properly. We can see this as follows.

The Helium Atom (cont.)

When adding equal angular momenta, the top m state of the highest l ladder is

$$|l = 2\,l_1, m = 2\,l_1\,\rangle = |l_1, l_1\,\rangle^{(1)} \otimes |l_1, l_1\,\rangle^{(2)} \tag{18.21}$$

It is already symmetric, meaning that it is an eigenstate of the particle exchange operator P_{12} with eigenvalue $+1$. The angular momentum ladder operators do not affect the state's properties under particle exchange because the ladder operator is symmetric under particle exchange:

$$L_\pm^{(1)\otimes(2)} = L_\pm^{(1)} \otimes I^{(2)} + I^{(1)} \otimes L_\pm^{(2)} \tag{18.22}$$

In detail, let's act with P_{12}:

$$P_{12} L_\pm^{(1)\otimes(2)} |l = 2\,l_1, m = 2\,l_1\,\rangle = \left(P_{12} L_\pm^{(1)\otimes(2)} P_{12}\right) \left(P_{12} |l = 2\,l_1, m = 2\,l_1\,\rangle\right) \tag{18.23}$$

$$= L_\pm^{(1)\otimes(2)} |l = 2\,l_1, m = 2\,l_1\,\rangle \tag{18.24}$$

where we have used $P_{12}^2 = I$. So, we see that $L_\pm^{(1)\otimes(2)} |l = 2\,l_1, m = 2\,l_1\,\rangle$ is also an eigenstate of the exchnange operator with eigenvalue $+1$. This of course continues all the way down the ladder.

The Helium Atom (cont.)

Now, consider the next lower total angular momentum, $l = 2 l_1 - 1$. We know that the top state of this ladder is obtained by orthogonality to the next-to-top state of the $l = 2 l_1$ ladder, so it is always

$$
\begin{aligned}
|\, l &= 2 l_1 - 1, m = 2 l_1 - 1 \,\rangle \\
&= (2^{-1/2}) \left(|\, l_1, l_1 \,\rangle^{(1)} \otimes |\, l_1, l_1 - 1 \,\rangle^{(2)} - |\, l_1, l_1 - 1 \,\rangle^{(1)} \otimes |\, l_1, l_1 \,\rangle^{(2)} \right)
\end{aligned} \quad (18.25)
$$

This state is clearly also an eigenstate of the particle exchange operator with eigenvalue -1. Again, the angular momentum ladder operators do not affect this property, so all states of the $l = 2 l_1 - 1$ ladder are antisymmetric under particle interchange.

I have so far been unable to prove in a generic way that this pattern continues, with the eigenvalue under particle exchange being $(-1)^l$, but it is almost assuredly true. Certainly, one can prove it explicitly for any specific case you want by just looking at the Clebsch-Gordan expansion. (I think a general proof will somehow connect the particle exchange operation to the parity operation.) The generic point is that one does not have to explicitly symmetrize when $n_1 = n_2$, $l_1 = l_2$ because the standard angular momentum addition rules already do it for you.

Let's talk a bit about the space that these symmetrized and antisymmetrized states live in for the case $n_1 \neq n_2$ or $l_1 \neq l_2$ because it is somewhat surprising. When adding angular momenta l_1 and l_2, one is used to seeing states that belong to the space

$$V^{(l_1)} \otimes V^{(l_2)} = V^{(l_1+l_2)}_{l_1 l_2} \oplus \cdots \oplus V^{(|l_1-l_2|)}_{l_1 l_2} \qquad (18.26)$$

where the superscripts indicate the total orbital angular momentum and the denominator indicates the individual orbital angular momenta that contributed. That is, the first state in the direct product is always l_1 and the second state is always l_2.

The Helium Atom (cont.)

Here, we are first constructing a bigger space because symmetrization/antisymmetrization requires that we also consider states with the orbital angular momenta exchanged, but then we restrict to a subspace thereof of eigenstates of the exchange operator. Let's consider the case $n_1 = n_2$ for notational brevity; one could generalize the following to $n_1 \neq n_2$ if one wanted. For this case, we first construct the larger space

$$
\left(V^{(l_1)} \otimes V^{(l_2)} \right) \oplus \left(V^{(l_2)} \otimes V^{(l_1)} \right)
$$

$$
= \left(V^{(l_1+l_2)}_{l_1 l_2} \oplus \cdots \oplus V^{(|l_1-l_2|)}_{l_1 l_2} \right) \oplus \left(V^{(l_1+l_2)}_{l_2 l_1} \oplus \cdots \oplus V^{(|l_1-l_2|)}_{l_2 l_1} \right) \tag{18.27}
$$

$$
= V^{(l_1+l_2)}_{l_1 l_2} \oplus V^{(l_1+l_2)}_{l_2 l_1} \oplus \cdots \oplus V^{(|l_1-l_2|)}_{l_1 l_2} \oplus V^{(|l_1-l_2|)}_{l_2 l_1} \tag{18.28}
$$

Then we restrict to

$$
\left(V^{(l_1)} \otimes V^{(l_2)} \right)_{\pm} = \left[V^{(l_1+l_2)}_{l_1 l_2} \oplus V^{(l_1+l_2)}_{l_2 l_1} \oplus \cdots \oplus V^{(|l_1-l_2|)}_{l_1 l_2} \oplus V^{(|l_1-l_2|)}_{l_2 l_1} \right]_{\pm} \tag{18.29}
$$

$\pm$ implies the subspace consisting of the appropriate eigenstates of the exchange operator. These are the kind of states we have constructed on the previous page. Note, that the Clebsch-Gordan coefficients don't care whether the ordering is $l_1, m_1; l_2, m_2$ or $l_2, m_2; l_1, m_1$, so we can use the same CG coefficient for both pieces of each term in the sum over m_1.

One can also see how this more complicated construction is unnecessary when $l_1 = l_2$ (and $n_1 = n_2$) because the space is just

$$V^{(l_1)} \otimes V^{(l_1)} = V_{l_1 l_1}^{(2 l_1)} \oplus \cdots \oplus V_{l_1 l_1}^{(0)} \tag{18.30}$$

The notation suggests it is already symmetrized.

Now, we put together the pairings of spatial and spin states that have the necessary overall antisymmetry under exchange:

$$
|(n_1, n_2), l, m, (l_1, l_2); S = 1, M_S = \pm 1 \rangle
$$
$$
= |(n_1, n_2), l, m, (l_1, l_2) \rangle_- \otimes \left| \frac{1}{2}, \pm\frac{1}{2}; \frac{1}{2}, \pm\frac{1}{2} \right\rangle_+ \quad (18.31)
$$

$$
|(n_1, n_2), l, m, (l_1, l_2); S = 1, M_S = 0 \rangle
$$
$$
= |(n_1, n_2), l, m, (l_1, l_2) \rangle_- \otimes \left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle_+ \quad (18.32)
$$

$$
|(n_1, n_2), l, m, (l_1, l_2); S = 0, M_S = 0 \rangle
$$
$$
= |(n_1, n_2), l, m, (l_1, l_2) \rangle_+ \otimes \left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle_- \quad (18.33)
$$

We note that the first two states do not exist when $n_1 = n_2$, $l_1 = l_2$, and l is even because the $-$ orbital states vanish, reflecting Pauli exclusion, and similarly the last state does not exist when $n_1 = n_2$, $l_1 = l_2$ and l is odd because, again, the $-$ orbital states vanish due to Pauli exclusion. Note also that, for the analysis we have done so far, which does not include any perturbations, the energies of these states depend only on n_1 and n_2, so the above states are all still degenerate in energy for a given n_1, n_2 pair.

If one blindly follows the hydrogen atom analysis, one might be inclined at this point to try to construct states that are eigenstates of the $(J^2)^{(1)\otimes(2)}$ and $J_z^{(1)\otimes(2)}$ operator so that one can proceed with including fine structure (relativistic) effects. This is not the right way to proceed because we know those effects are of order 5×10^{-5} in energy in the H atom, and they can only increase by a factor of order Z^2 for the He atom, so they are still very small. As we shall see, the perturbation due to electrostatic repulsion of the two electrons is far bigger, of order 10%-20%. So, clearly, we should include that effect first and not yet worry about fine structure.

The Helium Atom (cont.)

Electrostatic Repulsion Perturbation

Now, let us consider the following perturbing Hamiltonian:

$$\delta H_{es}^{(1)} = \frac{e^2}{R_{12}} \tag{18.34}$$

where $R_{12} = |\vec{R}_1 - \vec{R}_2|$ is the operator for the distance between the two electrons. (This operator is clearly only obviously defined in the position basis.) For a given state $|(n_1, n_2), l, m, (l_1, l_2); S, M\rangle$, the energy shift will be, as usual,

$$E^{(1)} = \langle (n_1, n_2), l, m, (l_1, l_2); S, M | \frac{e^2}{R_{12}} | (n_1, n_2), l, m, (l_1, l_2); S, M \rangle \tag{18.35}$$

$$= \int d^3 r_1 \, d^3 r_2 \, \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} \, |\psi_{(n_1, n_2), l, m, (l_1, l_2)}(\vec{r}_1, \vec{r}_2)|^2 \tag{18.36}$$

which is going to be, in general, a mess. The radial part of ψ is easy because it will consist of two terms, one being $R_{n_1 l_1}(r_1) \, R_{n_2 l_2}(r_2)$ and the other being $R_{n_2 l_2}(r_1) \, R_{n_1 l_1}(r_2)$. But the angular part will in general be terrible because of all the different m_1, m_2 pairs that contribute to a given m value. We are saved some effort by the fact that $\delta H_{es}^{(1)}$ is clearly a scalar operator, so we will be able to use the Wigner-Eckart theorem to simplify the calculation somewhat, but there will still be lots of terrible integrals.

The Helium Atom (cont.)

Ground State First-Order Energy Correction

Let us thus restrict ourselves to calculating the ground state energy shift, which is

$$
E_0^{(1)} = \langle n_1 = 1, n_2 = 1, l = 0, m = 0, l_1 = 0, l_2 = 0; S = 0, M = 0 \,|
$$
$$
\frac{e^2}{R_{12}} | n_1 = 1, n_2 = 1, l = 0, m = 0, l_1 = 0, l_2 = 0; S = 0, M = 0 \rangle
\tag{18.37}
$$

$$
= \int d^3 r_1 \, d^3 r_2 \, \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} R_{10}^2(r_1) \, R_{10}^2(r_2) |Y_0^0(\theta_1, \phi_1)|^2 |Y_0^0(\theta_2, \phi_2)|^2
$$
$$
\langle S = 0, M = 0 \,| S = 0, M = 0 \rangle
\tag{18.38}
$$

$$
= \int d^3 r_1 \, d^3 r_2 \, \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} \left[4 \left(\frac{Z}{a_0} \right)^3 e^{-2 Z r_1 / a_0} \right] \left[4 \left(\frac{Z}{a_0} \right)^3 e^{-2 Z r_2 / a_0} \right] \left(\frac{1}{4\pi} \right) \left(\frac{1}{4\pi} \right)
\tag{18.39}
$$

$$
= \frac{1}{\pi^2} \left(\frac{Z}{a_0} \right)^6 \int d^3 r_1 \, d^3 r_2 \, \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} e^{-2 Z (r_1 + r_2) / a_0}
\tag{18.40}
$$

where we have included the change to the hydrogenic states due to replacing the hydrogenic potential e^2/r by $Z\, e^2/r$ (Z is the atomic number); it effectively just reduces the Bohr radius of the atom by a factor of Z.

Doing this integral is a tedious exercise in the use of special function identities. The basic idea is to use the fact that

$$\frac{1}{|\vec{r}_1 - \vec{r}_2|} = \frac{1}{\left(r_1^2 + r_2^2 - 2\,r_1\,r_2\,\cos\theta_{12}\right)^{(1/2)}} \tag{18.41}$$

$$= \frac{1}{r_>} \sum_{l=0}^{\infty} \left(\frac{r_<}{r_>}\right)^l P_l(\cos\theta_{12}) \tag{18.42}$$

$$= \frac{1}{r_>} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left(\frac{r_<}{r_>}\right)^l \frac{4\,\pi}{2\,l+1} Y_l^{m*}(\theta_1, \phi_1)\,Y_l^m(\theta_2, \phi_2) \tag{18.43}$$

where

$$r_< \min(r_1, r_2) \qquad\qquad r_> = \max(r_1, r_2) \tag{18.44}$$

$$\cos\theta_{12} = \cos\theta_1 \cos\theta_2 + \sin\theta_1 \sin\theta_2 \cos(\phi_1 - \phi_2) \tag{18.45}$$

θ_{12} is the polar angle $\vec{r}_{12}$ makes with the z-axis. The second identity relating $P_l(\cos\theta_{12})$ to a sum over products of spherical harmonics is known as the addition theorem for spherical harmonics and follows immediately from rules for addition of angular momentum projected onto the position basis.

The Helium Atom (cont.)

The above relation decomposes the integral into independent integrals over $\vec{r}_1$ and $\vec{r}_2$, which can then be done using the orthogonality properties of spherical harmonics (remember, we have Y_0^0 in our integral because it is a constant) and standard integration rules to obtain

$$E^{(1)} = \frac{5}{8} \frac{Z e^2}{a_0} \tag{18.46}$$

Numerically, we have

$$E^{(0)} = -108.85 \text{ eV} \qquad E^{(0)} + E^{(1)} = -74.83 \text{ eV} \qquad E^{true} = -79.02 \text{ eV} \tag{18.47}$$

So, we can see that: a) the correction is large, 40%; b) within 5% of the true value; and c) the difference between the perturbation theory value and the true value is of order $E^{(0)}\delta^2$ where $\delta = |E^{(1)}/E^{(0)}|$. Clearly, we were well justified in ignoring fine-structure corrections at this point.

We note that the above technique for doing the integration of the matrix element will not work in general: we used the fact that Y_0^0 and $\left(Y_0^0\right)^2$ are both constants so that we could replace $(Y_0^0)^2$ by Y_0^0 in the integral over each electron's position, which then made the angular integrals trivial. That will not be possible in general.

The Helium Atom (cont.)

Variational Estimate

Given how large the first-order perturbation correction is, it is clear we need to go to second order in perturbation theory to get an accurate estimate. This of course becomes quite painful because of the large number of integrals to be evaluated. (Though, as we will explain below, these integrals will only connect the ground state $l = 0$ states to other $l = 0$ states due to the fact that the perturbation is a scalar when viewed as a spherical tensor.)

An alternate method is to obtain a variational estimate of the energy. Because the ground state has the electrons in the an antisymmetric spin state, there is no "exchange-effect" repulsion of the spatial wavefunctions and we can imagine that a relatively simple trial wavefunction will be sufficient. We recognize that, in the ground state, because the electron states must be spherically symmetric, the primary effect of the electrostatic repulsion will be to screen the charge of the nucleus (remember, the nucleus is much smaller than the Bohr radius!). So, we try hydrogen-like trial wavefunction with an adjustable parameter that can be interpreted as allowing for a reduced apparent nuclear charge:

$$\psi(\lambda, r_1, r_2) = \frac{1}{\pi} \left(\frac{\lambda}{a_0} \right)^3 \exp^{-\lambda \, r_1/a_0} \, \exp^{-\lambda \, r_2/a_0} \tag{18.48}$$

An important property of this trial wavefunction is that it respects the particle-exchange and spherical symmetry requirements on the state — we shall see the same requirement in the Hartree-Fock formalism for many-electron atoms.

As usual, we minimize

$$E[\psi] = \frac{\langle \psi \,|H|\psi \rangle}{\langle \psi \,|\psi \rangle} \tag{18.49}$$

where $H = H_{2e} + H_{es}^{(1)}$. Using similar techqniues to those used for the first-order perturbation theory calculation, the integral becomes

$$E(\lambda) = \left(\lambda^2 - 2\,Z\,\lambda + \frac{5}{8}\,\lambda \right) \frac{e^2}{a_0} \tag{18.50}$$

and minimization with respect to E yields

$$\lambda = Z - \frac{5}{16} \qquad\qquad E = -\left(Z - \frac{5}{16} \right)^2 \frac{e^2}{a_0} = -77.50 \text{ eV for } Z = 2 \tag{18.51}$$

which is about three times more accurate than the first-order perturbation calculation.

The Helium Atom (cont.)

Excited State First-Order Perturbation Corrections

Doing the above kind of first-order perturbation calculation for excited states will necessitate doing lots of unilluminating integrals. So let's instead see what kind of qualitative understanding we can obtain. We shall see that we can make some generic statements about the difficulty of doing corrections to arbitrary excited states, and we can fully determine the structure of the corrections for $(1s)(nl)$ states, in particular breaking of l and S degeneracy, though we won't calculate the specific correction values themselves.

To do this, we need to consider the matrix element of $\delta H_{es}^{(1)}$ between a pair of states. The first thing to note is that $\delta H_{es}^{(1)}$ is a scalar operator with respect to total orbital angular momentum l and is the identity with respect to total spin angular momentum S. Therefore, we may immediately conclude

$$\pm \langle (n_1', n_2'), l', m', (l_1', l_2'), S', M' | \delta H_{es}^{(1)} | (n_1, n_2), l, m, (l_1, l_2), S, M \rangle$$
$$= \pm \langle (n_1', n_2'), l, m, (l_1', l_2') | \delta H_{es}^{(1)} | (n_1, n_2), l, m, (l_1, l_2) \rangle_\pm \, \delta_{l'l}\delta_{m'm}\delta_{S'S}\delta_{M'M}$$

This results in two simplifications.

The Helium Atom (cont.)

▶ When doing the first-order energy correction, we need to consider the matrix elements of $\delta H_{es}^{(1)}$ among energy-degenerate levels. Any pair (n_1, n_2) has a unique energy up to $n_1 \leftrightarrow n_2$, and this latter degeneracy is already included by symmetrization/antisymmetrization. When (n_1, n_2) are fixed, we are left with a $8\,n_1^2 n_2^2$-degenerate subspace – there are $2\,n_1^2$ possible (l_1, m_1, s_1) values, $2\,n_2^2$ possible (l_2, m_2, s_2) values, and then we may exchange the order of the two states ($\times 2$ more). But, when we rewrite this space in the (l, m, S, M) basis, we see from the above that only states that have the same (l, m, S, M) values have nonzero matrix elements of $\delta H_{es}^{(1)}$. The perturbation is diagonal in the total spin space because it is the identity operator in the spin space. So, when we fix (n_1, n_2, l, m), we only need to worry about the degeneracy between the states $|(n_1, n_2), l, m, (l_1, l_2); S, M\rangle$ with (n_1, n_2) allowed to be exchanged and with (l_1, l_2) allowed to vary while still giving (l, m), but with the same (l, m, S, M). So degeneracy really only becomes an issue in the subspace consisting of all (l_1, l_2) pairs for the given (n_1, n_2) (and its exchange) that can yield the specified (l, m). The size of this degenerate subspace will depend on n_1, n_2, l, and m, so we won't try to say anything more about it. We could probably write down formula for the size of these problematic degenerate subspaces, but there is little value in it.

- ▶ If we consider first-order wavefunction corrections and higher-order energy and eigenstate corrections, the off-diagonal matrix elements even for non-degenerate energy levels now matter. What the above formula tells us is that the state $|(n_1, n_2), l, m, (l_1, l_2); S, M\rangle_\pm$ can only receive contributions from states $|(n_1', n_2'), l, m, (l_1', l_2'), S, M\rangle_\pm$ such that (l_1', l_2') can yield (l, m). Also, note that, because S must be the same, the exchange symmetry of the spatial parts of the two states being connected must be the same: the exchange symmetry of the spatial state is set by the S value.

Rather than trying to do the generic case with the above useful but still not trivializing simplifications, let us next note that the energy needed to put both electrons in $n > 1$ states is greater than the energy needed to remove one electron completely:

$$E_{(n_1, n_2)} = -Z\,[Ry]\left(\frac{1}{n_1^2} + \frac{1}{n_2^2}\right) \quad \Longrightarrow \quad E_{1,\infty} = -Z\,[Ry] \qquad E_{2,2} = -\frac{Z\,[Ry]}{2}$$
(18.52)

Thus, the states with both $n_1 \neq 1$ and $n_2 \neq 1$ are not of much practical importance — the atom would rather ionize and become a 1-electron atom. So we restrict ourselves to the $(1s)(nl)$ states (with $nl \neq 1s$ since we have already dealt with the ground state).

The allowed unperturbed $(1s)(nl)$ states are

$$|(1, n_2), l, m, (0, l); S = 1, M = \pm 1 \rangle = |(1, n_2), l, m, (0, l) \rangle_- \otimes \left| \frac{1}{2}, \pm\frac{1}{2}; \frac{1}{2}, \pm\frac{1}{2} \right\rangle_+$$
(18.53)

$$|(1, n_2), l, m, (0, l); S = 1, M = \pm 0 \rangle = |(1, n_2), l, m, (0, l) \rangle_- \otimes \left| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \right\rangle_+$$
(18.54)

$$|(1, n_2), l, m, (0, l); S = 1, M = \pm 1 \rangle = |(1, n_2), l, m, (0, l) \rangle_+ \otimes \left| \frac{1}{2}, \pm\frac{1}{2}; \frac{1}{2}, \pm\frac{1}{2} \right\rangle_-$$
(18.55)

We will thus only consider corrections to these states.

For the above $(1s)(nl)$ states, the total angular momentum is entirely determined by the second electron, and the Clebsch-Gordan coefficients are all unity (the unitary transformation from the uncoupled to the coupled basis is just the identity operator). So, *all* the possible off-diagonal elements of $\delta H_{es}^{(1)}$ are eliminated — when $l_1 = 0$, there is only one (l_1, l_2) correction that yields (l, m), and that is $(l_1, l_2) = (0, l)$ (and its exchange). Since there are now no non-zero off-diagonal matrix elements of $\delta H_{es}^{(1)}$, the standard perturbation theory applies.

The first-order electrostatic energy correction thus takes the following form:

$$E_{es}^{(1)} = \langle (1, n_2), l, m, (0, l), S, M \, | \delta H_{es}^{(1)} | (1, n_2), l, m, (0, l), S, M \rangle \tag{18.56}$$

$$= {}_{\mp} \langle (1, n_2), l, m, (0, l) \, | \delta H_{es}^{(1)} | (1, n_2), l, m, (0, l) \rangle_{\mp} \tag{18.57}$$

$$2\, E_{es}^{(1)} = \langle 1, n_2, l, m, 0, l \, | \delta H_{es}^{(1)} | 1, n_2, l, m, 0, l \rangle + \langle n_2, 1, l, m, l, 0 \, | \delta H_{es}^{(1)} | n_2, 1, l, m, l, 0 \rangle$$
$$\mp \langle 1, n_2, l, m, 0, l \, | \delta H_{es}^{(1)} | n_2, 1, l, m, l, 0 \rangle \mp \langle n_2, 1, l, m, l, 0 \, | \delta H_{es}^{(1)} | 1, n_2, l, m, 0, l \rangle \tag{18.58}$$

The $\mp$ depends on S: $-$ for $S = 1$ and $+$ for $S = 0$. The 2 arises in the second term because of the $2^{-1/2}$ normalization factors in the symmetrized/antisymmetrized state.

In the above expression, $\delta H_{es}^{(1)}$ is symmetric under interchange between the two particles, so the first and second terms are equal and the third and fourth terms are equal

$$E_{es}^{(1)} = \langle 1, n_2, l, m, 0, l \,|\delta H_{es}^{(1)}|\, 1, n_2, l, m, 0, l \rangle \mp \langle n_2, 1, l, m, l, 0 \,|\delta H_{es}^{(1)}|\, 1, n_2, l, m, 0, l \rangle \tag{18.59}$$

$$\equiv A_{nl} \mp B_{nl} \tag{18.60}$$

The A_{nl} term is called the *Coulomb interaction energy* because it is just the matrix element of the Coulomb interaction between the state and itself, while the B_{nl} terms is called the *exchange interaction energy* because it is the matrix element of the Coulomb interaction between the state and its exchanged version. The first one we expect classically, the second we do not.

The Helium Atom (cont.)

Thus, excited states of the form $(1s)(nl)$ are split into two groups:

$S = 1$	spin symmetric	space antisymmetric	$A_{nl} - B_{nl}$	triplet	orthohelium
$S = 0$	spin antisymmetric	space symmetric	$A_{nl} + B_{nl}$	singlet	parahelium

As indicated above, we refer to these two kinds of states as "singlet" and "triplet" states because of the spin degeneracy for each level. There is of course more degeneracy than that — the energy does not depend on m or M at all — but the spin degeneracy is what distinguishes the two kinds of states. Also, recall that the m and M degeneracies are not "problematic" in that the matrix elements of $\delta H_{es}^{(1)}$ between states with different m or M vanish — there is no problem of perturbation expansion coefficients diverging.

The singlet and triplet states acquire the same Coulomb correction but different exchange corrections because of the symmetry of the spatial state. One can show that both A and B are positive; this certainly makes sense for A because A really is just the electrostatic Coulomb energy between the two electrons, which must always be positive energy because the Coulomb interaction is repulsive. The sign is not so obvious for B. Regardless, the result is that the singlet is always higher energy than the triplet. This makes sense because the triplet state, by requiring spatial antisymmetry under exchange, makes the spatial wavefunction vanish when the two electrons are at the same location and thus reduces the size of the electrostatic repulsion, while the electrostatic repulsion does not care about the spin state at all.

The other interesting thing to note is the terms "parahelium" and "orthohelium". The reason for this nomenclature is that transitions between the singlet and triplet states are very unlikely, and so, once a helium atom is put in one or the other of these two kinds of states, it stays in that state for a long time; effectively, parahelium and orthohelium act like two different kinds of atoms that cannot interconvert easily. Any atom that is put in one variant quickly drops down to the lowest energy state of that variant — $(1s)(1s)$ for para and $(1s)(2s)$ for ortho. Interconversions between these two states are very unlikely because they have the same $l = 0$ and different S (the para state is $S = 0$ and the ortho state is $S = 1$), and the EM interaction, which only interacts with orbital angular momentum (to first order in the EM field strength) and is a spherical vector interaction having only components $|1, \pm 1\rangle$ (no longitudinally polarized photons because the photon is massless), only can cause $\Delta l = \pm 1$ transitions (to first order in the EM field strength).

To summarize, the first-order electrostatic repulsion for $(1s)(nl)$ excited states provides, in addition to a significant energy correction as it does for the ground state, breaking of both the l and S degeneracies, with $S = 1$ having lower energy than $S = 0$ for any given $(1s)(nl)$ state. The m and M degeneracies are still present.

Generic Comments on Electrostatic Second-Order Perturbation Corrections

As we have discussed, the first-order electrostatic correction is only good to 5-10%. One can certainly do higher-order corrections. If we restrict to considering $(1s)(nl)$ excited states, then we obtain simplifications just as we did for the first-order energy shift: a state $|(1, n), l, m, (0, l), S, M\rangle$ has non-zero matrix elements of $\delta H_{es}^{(1)}$ only with states of the form $|(1, n'), l, m, (0, l), S, M\rangle$ — as with the first-order energy shift, the ambiguity about the various (l_1, l_2) combinations that can yield a given (l, m) goes away. Moreover, we immediately see that the only parameters on which the higher-order electrostatic energy shifts can depend are n, l, and S. We could try to write out an expression of the same kind we obtained for the first-order energy shift, but there is little reason to. First, our first-order energy shift expressions only made explicit the dependence on S; we did not calculate the dependence on n, l. Second, the sign of the second-order energy shift is not very important as they probably won't be big enough to reverse the ordering of the single and triplet energies obtained at first-order. Furthermore, it becomes harder to determine the sign in a generic fashion because, because the $(E_n^{(0)} - E_{n'}^{(0)})$ factor will result in terms of both signs for $n > 1$ in the expansion. So we will leave the electrostatic correction as is.

The Helium Atom (cont.)

He Atom Fine Structure

In principle, we would like to calculate the effect of relativistic and spin-orbit effects in the helium atom. However, this is made more complicated by the fact that the electrostatic repulsion effect is so large, which will modify the eigenstates substantially. Clearly, if we begin with the unperturbed eigenstates, which are only correct to 30-40% accuracy, our estimate for the fine structure correctsion will only be good to a similar accuracy. So we probably need to do the electrostatic correction to at least second order for the fine structure corrections to be accurate to 5-10%. Moreover, we would have to do the electrostatic correction to many orders so that the magnitude of the error in the electrostatic correction is small compared to the size of the fine-structure corrections, which we know from the H atom are of order 10^{-5}.

With this in mind, it becomes clear that we really only should concern ourselves with the energy degeneracies that can be broken by the relativistic corrections and not worry about the detailed values.

We know from the hydrogen atom that the relativistic kinetic energy correction depends on (n, l); the same should happen here. Our electrostatic correction already depends on n, l, so the relativistic KE correction will provide no additional breaking of degeneracies. So we do not bother with it.

The spin-orbit correction is more interesting, though. In the hydrogen atom, it broke j degeneracy (though it left l degeneracy). In the helium atom, the electrostatic correction breaks both the l degeneracy and, because of the singlet-triplet splitting, the S degeneracy. We shall see that the spin-orbit correction will break some of the remaining m and M degeneracy by creating a perturbation that depends on j and thus will care about m and M.

The Helium Atom (cont.)

It is obvious that the spin-orbit Hamiltonian for the He atom should be

$$\delta H_{so}^{(1)} = \sum_{i=1}^{2} \left(\vec{L}^{(i,o)} \otimes I^{(i,s)} \otimes \prod_{k \neq i} I^{(k,o)} \otimes I^{(k,s)} \right) \cdot \left(I^{(i,o)} \otimes \vec{S}^{(i,s)} \otimes \prod_{k \neq i} I^{(k,o)} \otimes I^{(k,s)} \right) \tag{18.61}$$

$$= \sum_{i=1}^{2} \vec{L}^{(i)} \cdot \vec{S}^{(i)} \tag{18.62}$$

where i indicates the particle number, o indicates the "orbital" part of the single-particle Hilbert space, s indicates the "spin" part, and the latter expression is the conventional shorthand we will use. The expression obviously generalizes for more than 2 electrons. That is, we just add up the spin-orbit interactions for each electron.

There are lower-level interactions between the spin of one electron and the orbital angular momentum of another (basically, one electron's orbit generates a magnetic dipole field that the other electron's spin sees) and between the spins of two electrons (again, one electron's spin generates a magnetic field that the other electron's spin sees); they are smaller because these depend on the electrons relative motion rather than each one's motion with respect to the stationary nucleus.

The Helium Atom (cont.)

One can show, though we will not prove it here, that

$$\langle (n_1, n_2), l, m, (l_1, l_2), S, M | \vec{L}^{(i)} \cdot \vec{S}^{(i)} | (n_1, n_2), l, m, (l_1, l_2), S, M \rangle$$
$$= \alpha((n_1, n_2), l, S) \langle (n_1, n_2), l, m, (l_1, l_2), S, M | \vec{L} \cdot \vec{S} | (n_1, n_2), l, m, (l_1, l_2), S, M \rangle \tag{18.63}$$

where

$$\vec{L} = \sum_{i=1}^{2} \vec{L}^{(i,o)} \otimes I^{(i,s)} \otimes \otimes \prod_{k \neq i} I^{(k,o)} \otimes I^{(k,s)} = \sum_{i=1}^{2} \vec{L}^{(i)} \tag{18.64}$$

$$\vec{S} = \sum_{i=1}^{2} I^{(i,o)} \otimes \vec{S}^{(i,s)} \otimes \otimes \prod_{k \neq i} I^{(k,o)} \otimes I^{(k,s)} = \sum_{i=1}^{2} \vec{S}^{(i)} \tag{18.65}$$

are the total orbital and spin angular momentum operators. *This jump from the single-particle spin-orbit Hamiltonian to the total spin-orbit Hamiltonian is not justified in most textbooks.* A proof is available in Section 18.5 of *Quantum Mechanics* by L. Marchildon. Not surprisingly, the proof relies on the ever-useful Wigner-Eckart theorem.

The Helium Atom (cont.)

So, the spin-orbit Hamiltonian's matrix elements will simply be

$$\langle (n_1, n_2), l, m, (l_1, l_2), S, M | \delta H_{so}^{(1)} | (n_1, n_2), l, m, (l_1, l_2), S, M \rangle \tag{18.66}$$

$$= \alpha((n_1, n_2), l, S) \langle (n_1, n_2), l, m, (l_1, l_2), S, M | \vec{L} \cdot \vec{S} | (n_1, n_2), l, m, (l_1, l_2), S, M \rangle \tag{18.67}$$

$$= \frac{1}{2} \alpha((n_1, n_2), l, S) \langle (n_1, n_2), l, m, (l_1, l_2), S, M | \left[J^2 - L^2 - S^2 \right] | (n_1, n_2), l, m, (l_1, l_2), S, M \rangle \tag{18.68}$$

$$= \frac{1}{2} \alpha((n_1, n_2), l, S) \hbar^2$$

$$\langle (n_1, n_2), l, m, (l_1, l_2), S, M | \left[\frac{J^2}{\hbar^2} - l(l+1) - S(S+1) \right] | (n_1, n_2), l, m, (l_1, l_2), S, M \rangle \tag{18.69}$$

The Helium Atom (cont.)

Because of our previous motivation to look at the $(1s)(nl)$ states, only let's consider that case only here. There are two subcases to consider. For $S = 0$, $j = l$ and eigenstates of L^2 and S^2 are eigenstates of J^2 also. Also, because $j = l$ and $S = 0$, the matrix element will vanish and there will be no energy correction; as is probably obvious, $\left\langle \vec{L} \cdot \vec{S} \right\rangle = 0$ for a $S = 0$ state. That is, we have

$$\langle (1, n), l, m, (0, l), 0, 0 \, | \delta H_{so}^{(1)} | (1, n), l, m, (0, l), 0, 0 \, \rangle = 0 \tag{18.70}$$

For $S = 1$ and $l = 0$, we have $j = s$, and it is like the $l = 1$, $S = 0$ case above with no energy shift (the l term contributes 0 and the j term cancels the s term). No diagonalization is necessary.

For $S = 1$ and $l > 0$, we are now adding angular momenta l and $S = 1$. We obtain $j = l + 1, l, l - 1$. The energy shift result will only depend on the j value, so we ought to transform to a new coupled basis that consists of eigenstates of J^2 and J_z. In that basis, the matrix elements are

$$\langle (1, n), j, m_j, l, S = 1, (l_1 = 0, l_2 = l) | \delta H_{so}^{(1)} | (1, n), j, m_j, l, S = 1, (l_1 = 0, l_2 = l) \rangle$$
$$= \frac{1}{2} \alpha((1, n), l, S = 1) \hbar^2 [j(j+1) - l(l+1) - 2] \tag{18.71}$$

So, we obtain a splitting of the three different j values that are allowed for this l, and all three will be different from the vanishing $S = 0$ case:

$$
\begin{array}{lll}
j = l + 1 & j(j+1) - l(l+1) - 2 = 2l \\
j = l & j(j+1) - l(l+1) - 2 = -2 \\
j = l - 1 & j(j+1) - l(l+1) - 2 = -2l - 2
\end{array}
\tag{18.72}
$$

We provide a table of the spin-orbit correction splittings for $(1s)(1s)$, $(1s)(2s)$, and $(1s)(2p)$;

nl	n	$l = l_2$	spatial sym.	S	spin sym.	j	$\frac{\langle \vec{L} \cdot \vec{S} \rangle}{\alpha \hbar^2}$	spectr. term	para/ortho
$1s$	1	0	$+1$	0	-1	0	0	$^1S_0$	P
$2s$	2	0	$+1$	0	-1	0	0	$^1S_0$	P
	2	0	-1	1	$+1$	1	0	$^3S_1$	O
$2p$	2	1	$+1$	0	-1	1	0	$^3L_1$	P
	2	1	-1	1	$+1$	0	-2	$^3L_0$	O
						1	-1	$^3L_1$	O
						2	$+1$	$^3L_2$	O

The Helium Atom (cont.)

One thing that may be confusing is the use of the terms "singlet/triplet," "para/ortho," and the question of whether there is a spin-orbit correction or not and how it breaks the degeneracy. "Para" states are always singlets and vice versa, and "ortho" states are always triplets and vice versa; this nomenclature depends only on the total spin $S = 0$ or $S = 1$. The spin-orbit correction for para (singlet) states always vanishes because $S = 0$. The spin-orbit correction for $L = 0$ ortho (triplet) states vanishes because $L = 0$. There will be a spin-orbit correction for $L > 0$ ortho (triplet) states, and there will be three distinct values for it, but the splitting does not separate states of the spin triplet, by which we means states of different M. The spin-orbit coupling splits states of different j, and we know J_z and S_z do not commute (m_j and M are not both good quantum numbers).

Lecture 49:
Many-Electron Atoms and the Hartree-Fock Method

Revision Date: 2008/03/12

The Hartree-Fock Method

Overview

Once the number of electrons exceeds two, treating the electrostatic repulsion due to other electrons by direct perturbation theory becomes intractable — the number of integrals to calculate goes up as $n!$, and the approximation becomes ever worse. To some extent, when there are closed shells (*i.e.*, all the $2s$ states are filled, all the $2p$ states are filled, etc.), one can treat the system as consisting of only the valence electrons with a modified (screened) Coulomb potential, but even this begins to fail for very high Z atoms.

The generic method for dealing with this is the "self-consistent field" approximation, or Hartree-Fock approximation. The idea is to make a guess at the eigenstates using using antisymmetrized linear combinations of unperturbed eigenstates, calculate an average electrostatic repulsive potential (the "mean field" or "central field"), adjust the coefficients in the expansion based on the mean field, recalculate the mean field, and so on to iteratively solve for the eigenstates.

We obviously cannot do such a solution exactly, but we outline the procedure for the sake of exposing you to it; it is an important generic computational technique.

The Hartree-Fock Method (cont.)

The Hartree Method

Let us first describe the Hartree method, which determines a mean field for the electrostatic repulsion but does not account for the identical nature of the electrons.

The basic idea is to use the variational technique to obtain a differential equation that can be solved iteratively in the method diescribed above. Let's assume that the ground state of a system of N electrons in the presence of a nucleus of atomic charge Z is described by a direct product state of the form

$$|\psi\rangle = |\psi^{(1)}\rangle \otimes \cdots \otimes |\psi^{(N)}\rangle \tag{18.73}$$

Since we will be working entirely in the position basis here, we may write this more explicitly as

$$\psi(\vec{r}_1, \ldots, \vec{r}_N) = \psi^{(1)}(\vec{r}_1) \cdots \psi^{(N)}(\vec{r}_N) \tag{18.74}$$

The Hamiltonian for the system in the position basis is

$$H = \sum_i H^{(i)} + \frac{1}{2} \sum_{i \neq j} V_{ij}(|\vec{r}_i - \vec{r}_j|) = \sum_i \frac{P_i^2}{2\,\mu} - \frac{Z\,e^2}{r_i} + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\vec{r}_i - \vec{r}_j|} \tag{18.75}$$

The Hartree-Fock Method (cont.)

So, the energy functional is

$$E[\psi] = \frac{\langle \psi \,|H|\psi \,\rangle}{\langle \psi \,|\psi \,\rangle} \tag{18.76}$$

More explicitly, we have

$$
\begin{aligned}
E[\psi] = & \\
& \left[\sum_i \int d^3 r_i \, \psi^{(i)*}(\vec{r}_i) \left(\frac{P_i^2}{2\,\mu} - \frac{Z\,e^2}{r_i} \right) \psi^{(i)}(\vec{r}_i) \prod_{\substack{j=1 \\ j \neq i}}^N \int d^3 r_j \, |\psi^{(j)}(\vec{r}_j)|^2 \right. \\
& \left. + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^N \int d^3 r_i \int d^3 r_j \, |\psi^{(i)}(\vec{r}_i)|^2 |\psi^{(j)}(\vec{r}_j)|^2 \frac{e^2}{|\vec{r}_i - \vec{r}_j|} \right] \left[\prod_{i=1}^N \int d^3 r_j \, |\psi^{(i)}(\vec{r}_j)|^2 \right]^{-1}
\end{aligned}
\tag{18.77}
$$

As we did for the single-particle variational technique, we wish to consider the change in E, δE, that arises due to a variation of one of the single-particle wavefunctions $\psi^{(i)}(\vec{r}_i)$ away from its ground-state value by a deviation $\delta\psi^{(i)}(\vec{r}_i)$. We have

$$
E[\psi + \delta\psi^{(i)}(\vec{r}_i)] =
$$

$$
\left[\sum_i \int d^3 r_i \left[\psi^{(i)}(\vec{r}_i) + \delta\psi^{(i)}(\vec{r}_i) \right]^* \left(\frac{P_i^2}{2\,\mu} - \frac{Z\,e^2}{r_i} \right) \psi^{(i)}(\vec{r}_i) \prod_{\substack{j=1 \\ j\neq i}}^{N} \int d^3 r_j \, |\psi^{(j)}(\vec{r}_j)|^2 + H.C. \right.
$$

$$
+ \frac{1}{2} \sum_{\substack{i,j=1 \\ i\neq j}}^{N} \int d^3 r_i \int d^3 r_j \left[\psi^{(i)}(\vec{r}_i) + \delta\psi^{(i)}(\vec{r}_i) \right]^* \psi^{(i)}(\vec{r}_i) |\psi^{(j)}(\vec{r}_j)|^2 \frac{e^2}{|\vec{r}_i - \vec{r}_j|} + H.C. \Bigg]
$$

$$
\times \left[\int d^3 r_i \left[\psi^{(i)}(\vec{r}_i) + \delta\psi^{(i)}(\vec{r}_i) \right]^* \psi^{(i)}(\vec{r}_i) \prod_{\substack{j=1 \\ j\neq i}}^{N} \int d^3 r_j \, |\psi^{(j)}(\vec{r}_j)|^2 + H.C. \right]^{-1}
$$

$$
\tag{18.78}
$$

where $H.C.$ stands for "hermitian conjugate."

We move the denominator to the left side, subtract of $E[\psi]$ multiplied by its denominator, and define $\delta E = E[\psi + \delta\psi^{(i)}(\vec{r}_i)] - E[\psi]$ to obtain

$$
\delta E \left[\prod_{i=1}^{N} \int d^3 r_j \, |\psi^{(i)}(\vec{r}_j)|^2 \right]
$$

$$
+ E \left[\int d^3 r_i \left[\delta\psi^{(i)}(\vec{r}_i) \right]^* \psi^{(i)}(\vec{r}_i) \prod_{\substack{j=1 \\ j \neq i}}^{N} \int d^3 r_j \, |\psi^{(j)}(\vec{r}_j)|^2 + H.C. \right]
$$

$$
= \left[\sum_i \int d^3 r_i \left[\delta\psi^{(i)}(\vec{r}_i) \right]^* \left(\frac{P_i^2}{2\,\mu} - \frac{Z\,e^2}{r_i} \right) \psi^{(i)}(\vec{r}_i) \prod_{\substack{j=1 \\ j \neq i}}^{N} \int d^3 r_j \, |\psi^{(j)}(\vec{r}_j)|^2 + H.C. \right.
$$

$$
\left. + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \int d^3 r_i \int d^3 r_j \left[\delta\psi^{(i)}(\vec{r}_i) \right]^* \psi^{(i)}(\vec{r}_i) |\psi^{(j)}(\vec{r}_j)|^2 \frac{e^2}{|\vec{r}_i - \vec{r}_j|} + H.C. \right]
$$

The Hartree-Fock Method (cont.)

We set $\delta E = 0$ and take the variations $\delta \psi^{(i)}(\vec{r}_i)$ and $[\delta \psi^{(i)}(\vec{r}_i)]^*$ to be independent and arbitrary, which implies that the coefficient of each of these variations in the integrands must vanish at every point. We also use the fact that each individual particle's $\psi^{(i)}$ is normalized. We thus obtain

$$
\left(\frac{P_i^2}{2\,\mu} - \frac{Z\,e^2}{r_i} \right) \psi^{(i)}(\vec{r}_i) + \left(\frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \int d^3 r_j \, |\psi^{(i)}(\vec{r}_j)|^2 \frac{e^2}{|\vec{r}_i - \vec{r}_j|} \right) \psi^{(i)}(\vec{r}_i) = E\,\psi^{(i)}(\vec{r}_i)
$$

(18.79)

and its hermitian conjugate equation. We thus see that each particle obeys a single-particle Schrödinger Equation for a particle in a potential that is the sum of the attractive hydrogenic potential and a repulsive potential whose value is determined by the wavefunctions of all the other electrons. This latter potential is just the "mean field" due to the other electrons. We have N coupled differential equations for the N single-particle wavefunctions $\psi^{(i)}(\vec{r}_i)$ that make up $\psi(\vec{r}_1, \dots, \vec{r}_N)$.

The standard method for solving the above set of coupled equations is to employ an interative method, making initial guesses that might just be pure hydrogenic states, calculating the mean field term for each equation, solving each equation for a better approximation to the true wavefunctions, using these better approximations to calculate the mean field term, and so on. The process will converge.

The Hartree-Fock Method (cont.)

The Hartree-Fock Method

The above derivation does not take into account the fact that the electrons are identical and so the states must be antisymmetric under particle interchange. We need to include this equirement in the calculation.

At this point, it becomes absolutely necessary to incorporate spin, as this determines whether the spatial wavefunction needs to be symmetric or antisymmetric under particle interchange. This inclusion is not so complicated as it may seem, though, because the electrostatic potential is a scalar operator. Therefore, it commutes with the total orbital angular momentum operator $\vec{L}$, and therefore l, m are good quantum numbers. Furthermore, the electrostatic repulsive potential commutes with $\vec{S}$, so S, M are also good quantum numbers. So, we can construct states of definite l, m, S, M from hydrogenic states by:

1. Writing out non-interacting, non-identical single-electron hydrogenic states with definite l_i, m_i, $m_{s,i}$ (as well as n_i);
2. Constructing non-interacting, non-identical N-electron states of definite l, m (as well as $\{n_i\}$) from the single-electron states of definite l_i, m_i;
3. Constructing the symmetric and antisymmetric linear combinations of the above states to form symmetrized/antisymmetrized N-electron states of definite l, m and $\{n_i\}$, though symmetrization now implies that no one electron resides in a particular l_i, m_i, n_i;

The Hartree-Fock Method (cont.)

4. Constructing non-interacting, non-identical N-electron spin states of definite S, M;

5. Constructing the symmetric and antisymmetric linear combinations of the above states to form symmetrized/antisymmetrized N-electron spin states of definite S, M, though symmetrization now implies no one electron resides in a particular $m_{s,i}$. Of course all the electrons are still in $s_i = 1/2$ states.

6. Taking products of the spatial and spin pieces to construct states that are overall antisymmetric under exchange.

Now, one may apply the Hartree equations and method, requiring that the states used to solve the Hartree equations be of the above type, exchange-antisymmetrized eigenstates of L^2, L_z, S^2, and S_z with corresponding quantum numbers l, m, S, M. Restriction to these kinds of states enforces the identical-particle antisymmetry requirement as well as the scalar nature of the electrostatic repulsion operators.

No further closed-form results are possible at this point.

The Hartree-Fock Method (cont.)

Hund's Rule

We are able to state one generic result that follows from the antisymmetrization requirement of the Hartree-Fock method. Given N electrons, we are assured that the combination with total angular momentum $S = N/2$ will be symmetric under exchange because the topmost M for this value of S can only be built from the completely symmetric combination of the N electron spins. The spatial state that is partnered with this spin state must therefore be completely antisymmetric. Of all possible spatial states, this combination proves the smallest electrostatic repulsion correction because antisymmetrized states tend to yield maximum spatial separation of the electrons. Therefore, this state will have the lowest energy of all states with the same principal quantum numbers, just as the $S = 1$ states are always lower in energy than the $S = 0$ states of the helium atom. This rule that the highest total spin state is the lowest energy state is known as Hund's Rule and can be used to guess the ground-state electronic configuration and spectroscopic term for many-electron atoms.