Lecture Notes for "Topics in Spin"

Lead Instructor: Stephen Sekula
Contributions from: Tingting Cao, Kent Hornbostel, and Roberto Vega

April 10, 2014

Contents

1 Resources

The sources for this lecture series are:

- Halzen, F. and Martin, A. "Quarks and Leptons: An Introductory Course in Modern Particle Physics"
- McMahon, D. "Quantum Mechanics Demystified (a self-teaching guide)" (Hey! I am an experimentalist...)
- Merzbacher, E. "Quantum Mechanics (3rd Edition)"
- Sakurai, J. J. "Modern Quantum Mechanics"
- Shankar, R. "Basic Training in Mathematics: A Fitness Program for Science Students"
- Shankar, R. "Principles of Quantum Mechanics"
- Strang, G. "Linear Algebra and its Applications"
- Tomonaga, S. "The Story of Spin"
- Wu Ki Tung, "Group Theory in Physics"

You should also mine the bibliography at the end of these lecture notes for references to historical physics publications that were part of the long development of quantum mechanics and an understanding of spin angular momentum.

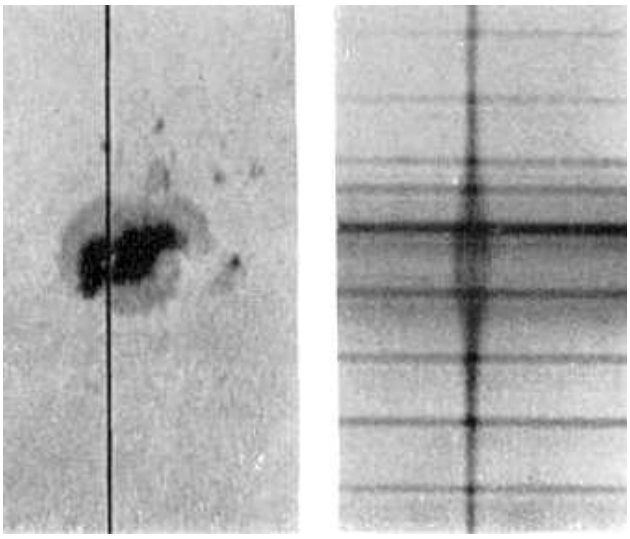
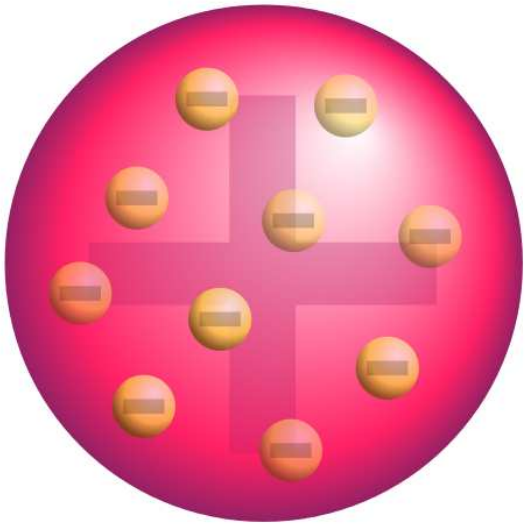


Figure 1: The Zeeman Effect, visible in a solar spectral line (left) that splits in the presence of a strong magnetic field gradient (right). Look at the bowed portion of the spectral line at the elevations where the sunspot regon is located. Sunspots are regions of high solar magnetic field gradient. This photo was taken in 1919, and is from Ref. [4].

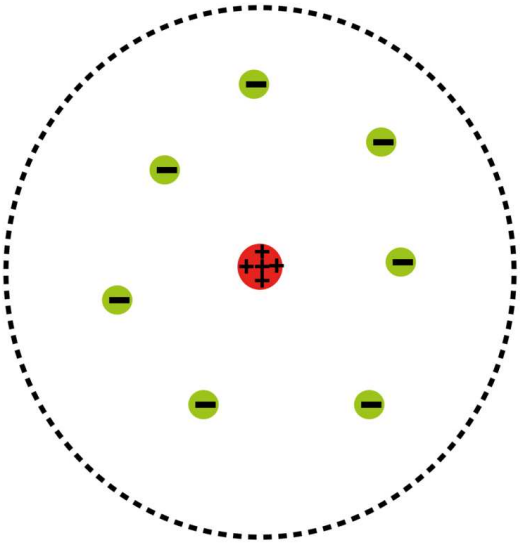
2 A very brief and over-simplified history of spin angular momentum

The development of quantum mechanics happened in parallel with a vast array of experimental observations regarding atoms. These observations were made in detail through the 19th century and into the early 20th century. It's useful to see how the model of the atom changed in response to experimental observations (data).

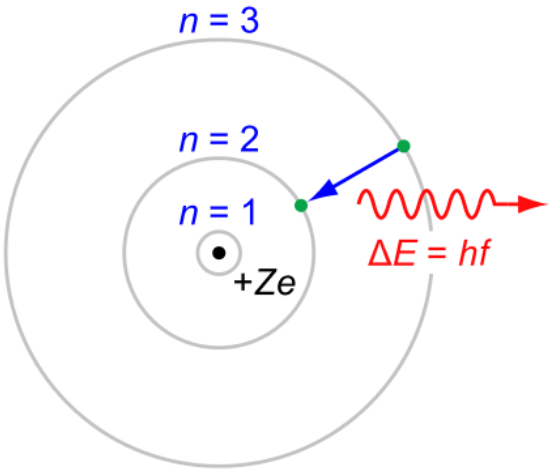
- In 1897, Zeeman reports that atomic spectral emissions will multiply (into doubles or triples) in the presence of a strong magnetic field ("The Zeeman Effect") [1, 2, 3]. The Zeeman Effect is illustrated in an astrophysics context in Fig. 1.
- The existence of atoms was finally and firmly established by Albert Einstein's 1905 paper regarding Brownian Motion[5]. This paper explained, using the atomic hypothesis (that matter is composed of fundamental building blocks), the observed phenomenon of Brownian Motion (the jittering of particles in a hot liquid), which other competing hypotheses for the nature of elements and compounds could not explain.
- Experiments on the electron, perhaps the most famous being those of J. J. Thompson[6, 7], established that the electron was not a wave nor an atom, but rather an independent particle (1897). In order to include the electron into the atomic model, Thompson proposed that the atom was composed of a continuous distribution of electrons and some positive charge elements so that their total electric charge was zero (neutral). This he proposed in 1904 (Fig. 2a). This exact picture of the distribution (sometimes described as the "plum pudding model") continued to shift absent experimental evidence for it or against it.
- In 1911, Ernest Rutherford performed his famous scattering experiments [11]and demonstrated that the atom was composed of a tightly-packed core nucleus of positive charge. The picture was then that the nucleus was at the center and a cloud of electrons surrounded the nucleus, forming the atom. The earlier "plum pudding"-style models were abandoned in favor of this picture. The new model is a "planetary model of the atom" - the electrons orbit a central nucleus (Fig. 2b). This model has a flaw. One could calculate that single electrons in such orbits would radiate energy, eventually leading to the collapse of their orbits and the dissolving of atoms. All atoms should be unstable.



(a) The "Plum Pudding" Model of the Atom.



(b) The Rutherford Model of the atom



(c) The Bohr-Rutherford Model of the atom

Figure 2: Cartoon representations of various ad hoc models of the atom, determined from experiments conducted from 1897-1922. The figures are from Refs. [8, 9, 10].

- The earlier observation (in the late 19th century) that the energy emitted from electric discharges of atoms was discrete and not continuous implied, in the Planetary Model, that electrons were forced into only very specific orbits, and could not occupy intermediate orbits. The problem of radiation, and these earlier observations of spectra interpreted in the Planetary Model, led to the Bohr-Rutherford Model of the Atom in 1913[12, 13, 14, 15, 16] (Fig. 2c). This was the next logical step beyond the planetary model also consistent with the data (lack of observation of general atomic instability and discrete atomic emission spectra). Electrons surround the nucleus, but only in well-defined orbits and not anywhere in-between.

Atoms in the Bohr Model then have discrete orbits, leading to discrete energies emitted during transitions between the orbits:

$$\Delta E = E_2 - E_1 = hf$$

where h is Planck's Constant ($h = 6.626 \times 10^{-34}\,\mathrm{J \cdot s}$) and f is the frequency of the emitted light. Discrete orbits have specific orbital angular momentum associated with them, and that, too, will be discrete:

$$L = n\hbar$$

where n is the "Principle Quantum Number" assigning the electron to a specific orbital level. This notation was introduced by Bohr.

Arnold Sommerfeld and Peter Debye independently proposed an enhancement of the Bohr Model[17] and were the first to formally propose the discretization of a *component* of total orbital angular momentum ("space quantization"). They proposed that orbits could deviate from circular (as was the assumption of the Bohr Model), stretching along an axis as ellipses. This required introducing not just a principle quantum number, n, to describe the orbital level, but also a new quantum number, k, to describe the shape of the orbit (deviation from circular). Sommerfeld proposed that k can only take positive integer values (preserving the discretization of orbital angular momentum) but the stretch of the orbit could vary; however, the stretch could also only vary in a discrete manner. This was described by a quantum number m, such that $-k \leq m \leq k$. Sommerfeld thus predicted that orbital angular momentum should be quantized along a specific axis, a prediction that was tested later by the Stern-Gerlach Experiment (1922) [18] and verified (for a review of the Stern-Gerlach Experiment, c.f. Ref. [19]). Space quantization failed to illuminate the Zeeman Effect, however.

The Stern-Gerlach Experiment, demonstrated space quantization but it took time to understand exactly the results. It wasn't until the postulation of electron "spin angular momentum" that the specific reasons for the observed effect could be completely understood. This wasn't until 1927. The concept of a two-valued internal degree of freedom present in the electron - what we call "spin" - was thanks to Wolfgang Pauli in 1924.

The early-mid 1920s mark the end of the "old quantum mechanics," which effectively was a series of ad hoc models built in response to experimental observations, and the maturation of the development of formal quantum mechanics based on various principles, including energy conservation and the observed wave nature of matter and light. Formal quantum mechanics allowed for an exact mathematical model of the atom to be built, and that model reproduced exactly many of the earlier observed atomic phenomena while predicting new ones. Special relativity was incorporated into formal quantum mechanics by Paul Dirac in 1928 (c.f. [20]).

In these lectures, I will walk a narrow path through the subject of spin. We will roughly follow this trajectory:

- A review of the most salient and basic mathematical tools needed to make progress
- The application of the tools to orbital angular momentum as a preparation and refresher on the subject
- The introduction, by hand, of spin angular momentum into non-relativistic quantum mechanics
- The handling of simple multi-particle systems and the addition of individual angular momentum states to obtain a single, total angular momentum state

- Spin in relativistic quantum mechanics
- Spin and the identification of new particles, with a focus on the Higgs Boson and the theoretical and experimental realities of measuring its spin quantum number

3 Basic Mathematical Tools

3.1 Basic Information About Complex Numbers

Consider a complex number, z, which can be written as the sum of real and imaginary parts:

$$z = \mathcal{R}(z) + i\,\mathcal{I}(z) = x + iy.$$

The complex conjugate of this number can be written as:

$$z^* = x - iy$$

A few interesting relationships can be immediately derived:

$$\begin{aligned} \mathcal{R}(z) &= \frac{z + z^*}{2} \\ \mathcal{I}(z) &= \frac{z - z^*}{2i}. \end{aligned}$$

Using this, we can also write the magnitude of the complex number (which must be a real number):

$$\begin{aligned} |z|^2 &\equiv zz^* \\ &= \mathcal{R}(z)^2 + \mathcal{I}(z)^2 \\ &\geq \mathcal{I}(z)^2 = \left(\frac{z - z^*}{2i}\right)^2. \end{aligned}$$

3.2 Vectors, Vector Spaces, and the Dirac Notation

3.2.1 Crutch: vectors in space and time

When we learn vectors, we learn about space and then space-time vectors:

$$\begin{aligned} \vec{x} &= (x, y, z) \\ X &= (x, y, z, ct) = (\vec{x}, ct). \end{aligned}$$

3.2.2 Generalization: vectors in the Dirac Notation

However, vectors are not restricted to only being columns or rows of numbers. They can be collections of any type of mathematical object - functions, matrices, etc. In order to generalize the concept of a vector, we introduce the Dirac Notation. A column-vector is referred to as a *ket*, and is denoted:

$$|x\rangle = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

while a row-vector is referred to as a *bra* (bra-ket . . . get it? I know. Terrible. It annoyed me when I first learned this notation.) and is denoted:

$$\langle x| = \begin{bmatrix} x & y & z \end{bmatrix}.$$

I have above used real numbers in the construction of a vector; however, we should not limit ourselves to real numbers, but generalize to complex numbers. A set of vectors is said to be constructed over the "field of real numbers" if all elements of the vectors in the set are real-valued; likewise, a set is constructed over the "field of imaginary" or the "field of complex" numbers if elements of the vectors in the set are constructed from imaginary or complex numbers, respectively.

It is important to note that the most general form of the row-vector is:

$$\langle x| = \begin{bmatrix} x^* & y^* & z^* \end{bmatrix} = \begin{bmatrix} x^* \\ y^* \\ z^* \end{bmatrix}^T = (|x\rangle)^{*T} .$$

That is, the *bra* is simply the transpose of the complex-conjugate of the *ket*.

3.2.3 Vector Spaces

A *Vector Space,*$\mathbb{V}$*,* is defined according to these properties:

- There is a definite rule for forming the *sum* of two vectors, denoted $|V\rangle + |W\rangle$.
- There is a definite rule for multiplication by scalars a,b,... denoted $a\,|V\rangle$with the following features:
  - The result of these operations (addition, and multiplication by scalars) results in another vector in the same space, a feature called "*closure*": $|V\rangle + |W\rangle \in \mathbb{V}$.
  - Scalar multiplication is *distributive*: $(a+b)\,|V\rangle = a\,|V\rangle + b\,|V\rangle$.
  - Scalar multiplication is *associative*: $a\,(b\,|V\rangle) = (ab)\,|V\rangle$
  - Addition is *commutative* (independent of the order of addition): $|V\rangle + |W\rangle = |W\rangle + |V\rangle$.
  - Addition is *associative*: $(|V\rangle + |W\rangle) + |Z\rangle = |V\rangle + (|W\rangle + |Z\rangle)$.
  - There exists a *null vector* obeying: $|V\rangle + |0\rangle = |V\rangle$.
  - For every vector there exists an inverse under addition, such that: $|V\rangle + |-V\rangle = |0\rangle$.

The scalars involved above are called the *field* over which the vector space is defined, as mentioned earlier. As Shankar says in his text on quantum mechanics, it's fairly easy to remember all of these rules... "do what comes naturally."

We will deal only with linear vector spaces - that is, those where any vector in the space can be written as a simple linear sum of other vectors in the space:

$$|W\rangle = \sum_{i=1}^{n} a_i\,|V_i\rangle\,.$$

3.2.4 Linear Independence and Basis Vectors

Here, we can define an important feature of some subset of vectors in the space. Consider this sum:

$$\sum_{i=1}^{n} a_i\,|V_i\rangle = |0\rangle$$

A set of vectors, $|V_i\rangle$, is said to be *linearly independent* if and only if the above sum is achieved by setting all $a_i = 0$. That is, the set is linearly indepenent if there is no combination of multiplicative scalars that leads to a null sum, other than when the scalars are themselves null.

A set of linearly independent vectors whose sum can be used to obtain any other vector in the space is referred to as *the basis of the vector space.*

3.2.5 The (Inner) Scalar Product

Consider a vector $|\psi\rangle$. The scalar product is that product which returns a number in the field on which the vector space is defined. When the scalar product is computed between a vector and itself, we refer to this as the "magnitude" of the vector.

The general scalar product is called the "inner product" and is denoted simply:

$$\langle W|V\rangle \in \mathbb{R}$$

The magnitude (length) of a vector is then given by:

$$\begin{aligned} |V|^2 &= \langle V|V\rangle \\ |V| &= \sqrt{\langle V|V\rangle}. \end{aligned}$$

We demand that this product obey the following axioms:

- Skew-symmetry: $\langle V|W\rangle = \langle W|V\rangle^*$
- Positive semidefiniteness: $\langle V|V\rangle \geq 0$ and is exactly 0 iff $|V\rangle = |0\rangle$.
- Linearity in ket: $\langle V|\left(a\,|W\rangle + b\,|Z\rangle\right) = \langle V|(aW + bZ\rangle = a\langle V|W\rangle + b\langle V|Z\rangle$.

Such a vector space with this product defined is called an *inner product space.*

- Two vectors are *orthogonal* if their inner product vanishes
- $|V|$ is referred to as the "norm" or "length" of the vector, $|V\rangle$.
- A set of basis vectors all of unit norm is referred to as an "orthonormal basis."

3.2.6 Expansion of vectors in an orthonormal basis

If we have identified a set of orthonormal basis vectors, $|i\rangle$, any vector in the space can be written as:

$$|V\rangle = \sum_{i=1}^{n} v_i\,|i\rangle\,.$$

To find the j^{th} component of this vector, we compute:

$$\begin{aligned} \langle j|V\rangle &= \sum_{i=1}^{n} v_i\langle j|i\rangle \\ &= \sum_{i=1}^{n} v_i\delta_{ij} \end{aligned}$$

where δ_{ij} is the *Kronecker Delta* which satisfies $\delta_{ii} = 1$ and $\delta_{ij} = 0$ for $i \neq j$. This merely yields:

$$\langle j|V\rangle = v_j.$$

Using this result, we can write the vector as an expansion in the basis vectors as follows:

$$|V\rangle = \sum_{i=1}^{n} |i\rangle\,\langle i|V\rangle.$$

In other words, a general vector may be written as the sum of basis vectors, where each basis vector is unit length and multiplied by a coefficient given by $\langle i|V\rangle$.

3.2.7 Finding an Orthonormal Basis - the Gram-Schmidt Theorem

We are often interested in finding the orthonormal basis vectors of a vector space. This can be done using the Gram-Schmidt procedure. Given a set of vectors that define a basis set (they must only be non-parallel), we can obtain an *orthonormal basis set* as follows:

1. Take one of the vectors, and merely rescale it by its own length so it becomes a unit vector. This yields the first normal basis vector.
2. Subtract from the second vector its projection along the first, leaving behind only the part perpendicular to the first. Rescale this piece by its own length, yielding a second normal vector orthogonal to the first.
3. Repeat this procedure; for each additional vector in the original basis, subtract from it the projections of its length along the other orthonormal vectors, and rescale the result by its length.

Writing out the steps in Dirac Notation:

1. The first unit-length vector, $|1\rangle$, is obtained from one of your basis vectors, $|I\rangle$, as follows:
$$|1\rangle = \frac{|I\rangle}{\sqrt{\langle I|I\rangle}}.$$
2. The second unit-length vector, is obtained by first doing this:
$$|2'\rangle = |II\rangle - |1\rangle\langle 1|II\rangle$$
and then finally by:
$$|2\rangle = \frac{|2'\rangle}{\sqrt{\langle 2'|2'\rangle}}.$$
3. The third is obtained via:
$$|3'\rangle = |III\rangle - |1\rangle\langle 1|III\rangle - |2\rangle\langle 2|III\rangle$$
followed by
$$|3\rangle = \frac{|3'\rangle}{\sqrt{\langle 3'|3'\rangle}}.$$
4. Rinse and repeat...

3.2.8 The Schwartz Inequality

This mathematical statement - the Schwartz Inequality - is essential in formulating a generic version of the Heisenberg Uncertainty Principle. It states that:

$$|\langle V|W\rangle| \leq |V||W|.$$

You can go and prove it, if you like; what you have learned from the earlier parts of the lecture should allow you to demonstrate this. *(HINT: you need to employ the axiom of positive semi-definiteness, $\langle Z|Z\rangle \geq 0$.)*

When two vectors are orthogonal, their inner product vanishes and this inequality is maximally true (0 is the smallest, positive, real number you can obtain from this product!). When two vectors are exactly parallel, then this is an exact equality. Non-parallel, non-orthogonal vectors lie in between.

In-Class Exercise: Practice with Vectors in a Matrix Representation
Consider the following vectors in a matrix representation:

$$|I\rangle = \begin{bmatrix} 1 \\ 5 \end{bmatrix}, \; |II\rangle = \begin{bmatrix} 7 \\ 0 \end{bmatrix}$$

1. Compute the complex-conjugate transposes of these vectors (the *bras* that correspond to these *kets*). This is referred to as the *adjoint* of the vector.
2. Calculate the length of each of these vectors.
3. Demonstrate that these vectors are non-parallel.
4. Using the Gram-Schmidt Theorem, create from these vectors an *orthonormal basis vector set.*

3.3 Linear Operators

An *operator* is any mathematical object that acts on a vector in the space, $\mathbb{V}$, and returns another vector in the space:

$$A\,|V\rangle = |W\rangle$$

or

$$\langle W|\,A = |Z\rangle$$

In the matrix representation, this would be represented by $n \times n$ matrices if the kets (bras) are represented by $n \times 1$ column matrices ($1 \times n$ row matrices).

We are concerned with *linear operators* that obey these rules:

$$\begin{aligned} A\alpha\,|V\rangle &= \alpha A\,|V\rangle \\ A\left(\alpha\,|V\rangle + \beta\,|W\rangle\right) &= \alpha A\,|V\rangle + \beta A\,|W\rangle \\ \langle V|\,\alpha A &= \langle V|\,A\alpha \\ \left(\langle V|\,\alpha + \langle W|\,\beta\right) A &= \alpha\,\langle V|\,A + \beta\,\langle W|\,A \end{aligned}$$

The simplest operator is *the identity operator*, which leaves the vector alone:

$$I\,|V\rangle = |V\rangle\,.$$

It satisfies this behavior for all kets and all bras.

3.3.1 Properties of Linear Operators

An operator, A, has an *inverse*, A^{-1}, if the following equation is satisfied:

$$AA^{-1} = I.$$

In general, the inverse is found by:

$$A^{-1} = \frac{A_C^T}{\det(A)}$$

where A_C^T is the *co-factor matrix* and det(A) is the *determinant.* It is good to here consult a devoted textbook on linear algebra.

A matrix is *Unitary* if

$$A^{*T}A \equiv A^\dagger A = I.$$

The operation of the *conjugate transpose* (A^{*T}) is known as determining the adjoint matrix, $A^\dagger$. In this very specific case, the inverse of the operator is its adjoint.

A matrix is *Hermitian* if it is its own self-adjoint; that is, if $A = A^\dagger$. As a result of this, the diagonal elements of a Hermitian matrix MUST be real numbers and thus $det(A)$ is a real number.

A matrix, A, is *positive-definite* if

$$\mathcal{R}(\langle x| A |x\rangle) > 0 \text{ for all } |x\rangle\,,\ x_i \in \mathcal{C}$$

(the real part of complex number that results from determining the projection of $A|x\rangle$ on the original $|x\rangle$ is positive, for all vectors formed on the field of complex numbers, x_i). For a Hermitian Matrix, the necessary (but not sufficient) conditions for being positive-definite are:

- $a_{ii} > 0$ for all i
- $a_{ii} + a_{jj} > 2|\mathcal{R}[a_{ij}]|$ for $i \neq j$
- The element with the largest modulus must lie on the main diagonal
- $det(A) > 0$

3.3.2 Matrix Elements of Linear Operators

The action of a linear operator can be fully specified by its actions on the basis vectors of a vector space:

$$A|i\rangle = |i'\rangle\,.$$

We can then write the action of the operator on any vector in the space:

$$A|V\rangle = A\sum_{i=1}^{n} v_i |i\rangle = \sum_{i=1}^{n} v_i A|i\rangle = \sum_{i=1}^{n} v_i |i'\rangle\,.$$

If we then take the inner product of this formula with another basis vector, $|j\rangle$,

$$\langle j|i'\rangle = \langle j| A |i\rangle \equiv A_{ij}$$

and we say that this is the i,j element of the operator in the matrix representation (the "matrix elements"). The components of a vector $|V'\rangle$ can then be expressed:

$$\begin{aligned} v'_i &= \langle i|V'\rangle = \langle i|A|V\rangle = \langle i|A\left(\sum_{j=1}^{n} v_j |j\rangle\right) \\ &= \sum_{j=1}^{n} v_j \langle i|A|j\rangle \\ &= \sum_{j=1}^{n} A_{ij} v_j. \end{aligned}$$

We can then imagine schematically how to determine the matrix elements of the operator from the basis vectors:

$$\begin{bmatrix} v'_1 \\ v'_2 \\ . \\ . \\ . \\ v'_n \end{bmatrix} = \begin{bmatrix} \langle 1|A|1\rangle & \langle 1|A|2\rangle & . & . & . & \langle 1|A|n\rangle \\ \langle 2|A|1\rangle & & & & & \\ & . & & & & \\ & . & & & & \\ & . & & & & \\ \langle n|A|1\rangle & & & & & \langle n|A|n\rangle \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ . \\ . \\ . \\ v_n \end{bmatrix}.$$

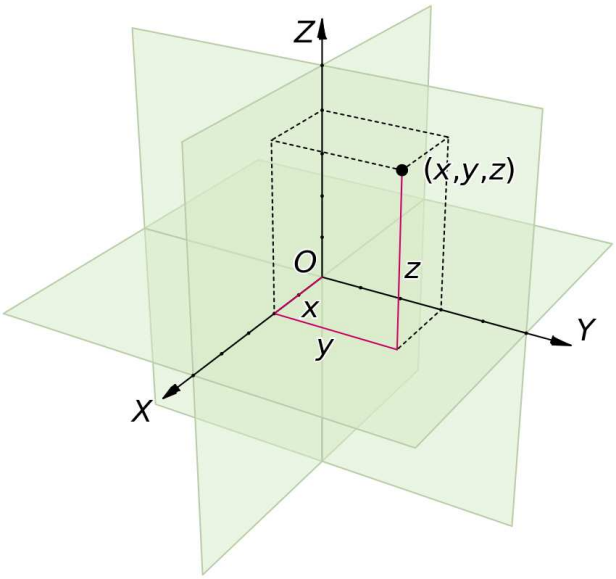


Figure 3: A 3-D Cartesian Coordinate system. This will help you to visualize rotations of the axes about the x-axis. From Ref. [21].

In-Class Exercise: The Rotation Operator, R
Consider an operator that executes rotations by $90°(\pi/2$ radians) around the unit vector, $\vec{i}$, in a Cartesian Coordinate System (Fig. 3) where the x-direction is denoted by $\vec{i}$, the y-direction by $\vec{j}$, and the z-direction by $\vec{k}$. The actions of the rotation operator are given as follows on the orthonormal basis vectors:

$$\begin{aligned} R(\frac{1}{2}\pi\vec{i})\,|i\rangle &= |i\rangle \\ R(\frac{1}{2}\pi\vec{i})\,|j\rangle &= |k\rangle \\ R(\frac{1}{2}\pi\vec{i})\,|k\rangle &= -\,|j\rangle \end{aligned}$$

Determine the matrix elements of this rotation operator.

3.3.3 Rotation and Commutation

The concept of a *rotation operator*, one that changes basis vectors into one another, is introduced in the in-class exercise above. In general, we are interested in understanding whether or not operators *commute* with one another - that is, whether or not the result of two operators acting on a basis vector depends on the order in which they are applied.

Consider the above rotation operator, and its equivalent for rotations about the y-axis:

$$\begin{aligned} R(\frac{1}{2}\pi\vec{j})\,|i\rangle &= -\,|k\rangle \\ R(\frac{1}{2}\pi\vec{j})\,|j\rangle &= |j\rangle \\ iR(\frac{1}{2}\pi\vec{j})\,|k\rangle &= |i\rangle \end{aligned}$$

Perform first a rotation of the y-axis around the x-axis, then rotate the result around the y-axis:

$$R(\frac{1}{2}\pi\vec{j})R(\frac{1}{2}\pi\vec{i})\,|j\rangle = R(\frac{1}{2}\pi\vec{j})\,|k\rangle$$

$$= |i\rangle .$$

Compare that to first rotating the y-axis around the y-axis, then rotating the result around the x-axis:

$$\begin{aligned} R(\frac{1}{2}\pi\vec{i})R(\frac{1}{2}\pi\vec{j})\,|j\rangle &= R(\frac{1}{2}\pi\vec{i})\,|j\rangle \\ &= |k\rangle . \end{aligned}$$

We see that we DO NOT GET THE SAME RESULT. The result is dependent on the order of operator application to the original state. This then let's us say that:

$$R(\frac{1}{2}\pi\vec{i})R(\frac{1}{2}\pi\vec{j}) \neq R(\frac{1}{2}\pi\vec{j})R(\frac{1}{2}\pi\vec{i})$$

or, similarly, that

$$R(\frac{1}{2}\pi\vec{i})R(\frac{1}{2}\pi\vec{j}) - R(\frac{1}{2}\pi\vec{j})R(\frac{1}{2}\pi\vec{i}) \neq 0$$

which is to say that these operators do not *commute* with one another. The commutation relation is written as:

$$[R(\frac{1}{2}\pi\vec{i}), R(\frac{1}{2}\pi\vec{j})] \equiv R(\frac{1}{2}\pi\vec{i})R(\frac{1}{2}\pi\vec{j}) - R(\frac{1}{2}\pi\vec{j})R(\frac{1}{2}\pi\vec{i}) \neq 0.$$

3.3.4 Eigenvalues and Eigenvectors

In general, an operator's action on a vector is some very complex return of rotation and other effects. However, there are a few privileged vectors that, when acted upon by an operator, are merely scaled by a number. These are called *eigenvectors* and the scaling factors are called *eigenvectors*.

The importance of this particular behavior cannot be understated, though it seems like a very simple thing at first. In equation form:

$$A\,|V\rangle = a\,|V\rangle .$$

We can write this in a suggestive equation form:

$$(A - aI)\,|V\rangle = |0\rangle$$

and then we see that the eigenvectors are given by:

$$|V\rangle = (A - aI)^{-1}\,|0\rangle .$$

There is a trivial solution . . . $|V\rangle = |0\rangle$. We're not interested in that. We are interested instead in the other case - where we consider the properties of $(A - aI)^{-1}$. The only way for a general matrix of this form to satisfy the above equation is if the determinant of the matrix vanishes, since:

$$(A - aI)^{-1} = \frac{(A - aI)^T_C}{det(A - aI)}$$

where C refers to the "co-factor matrix" of the original matrix. Consult a textbook on linear algebra for a deeper discussion of this. We only need one feature from the above equation: the only way for the eigenvector equation to be satisfied is if the determinant vanishes, since the cofactor matrix will be finite if the original matrix is finite. So we know that the eigenvector equation imposes:

$$det(A - aI) = 0.$$

The trick is finding these vectors. For instance, consider the two-dimensional rotation matrix that rotates $x \to y$ and $y \to -x$:

$$R_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

The eigenvalue equation is:

$$(R_2 - mI)\,|V\rangle = |0\rangle$$

From this, we know that:

$$\begin{aligned} det(R_2 - mI) &= 0 \\ det\begin{pmatrix} 0-m & 1 \\ -1 & 0-m \end{pmatrix} &= 0 \\ (0-m)^2 + 1^2 &= 0 \\ m^2 + 1 &= 0 \\ m &= \pm i. \end{aligned}$$

We now have the eigenvalues. To find the eigenvectors:

$$\begin{aligned} R_2\,|V\rangle &= +i\,|V\rangle \\ R_2\,|W\rangle &= -i\,|W\rangle \end{aligned}$$

Considering the first equation:

$$\begin{aligned} R_2\,|V\rangle &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = +i\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \\ \begin{matrix} v_2 \\ -iv_1 \end{matrix} &= \begin{matrix} iv_1 \\ iv_2 \end{matrix} \end{aligned}$$

We get a set of equations from which we can solve for the components - but many possible numbers will work here. What do we choose? Well, we also learn from the eigenvalue equation the following useful tidbit:

$$bA\,|V\rangle = ba\,|V\rangle = a(b\,|V\rangle) = a(A\,|V\rangle)$$

and any vector that is an eigenvector, when scaled, is *ALSO* an eigenvector. We can pick any vector we like that satisfies the above constraint equation, so let's choose a unit vector:

$$|V\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}.$$

The second eigenvector is then given by:

$$\begin{aligned} R_2\,|W\rangle &= -i\,|W\rangle \\ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} &= -i\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \\ \begin{matrix} w_2 \\ -w_1 \end{matrix} &= \begin{matrix} -iw_1 \\ -iw_2 \end{matrix} \end{aligned}$$

We can then write down a unit vector that satisfies these constraints:

$$|W\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}.$$

3.3.5 The Diagonal Form of a Matrix

The eigenvalue equation immediately let's us recognize the solution to a specific problem - what is the matrix, S, that transforms A such that the resulting matrix is entirely diagonal (only elements on the main diagonal)? This is represented by the equation:

$$S^{-1}AS = \Lambda$$

This is satisfied by constructing the matrix S from the eigenvectors, where each column is an eigenvector of A. If that is done, then we see that:

$$\begin{aligned} AS &= A\left(\, |V_1\rangle \quad . \quad . \quad . \quad |V_n\rangle \, \right) = \left(\, a_1 \, |V_1\rangle \quad . \quad . \quad . \quad a_n \, |V_n\rangle \, \right) \\ &= \left(\, |V_1\rangle \quad . \quad . \quad . \quad |V_n\rangle \, \right) \begin{pmatrix} a_1 & & & & \\ & . & & & \\ & & . & & \\ & & & . & \\ & & & & a_n \end{pmatrix} = S\Lambda \end{aligned}$$

and therefore:

$$S^{-1}(AS) = S^{-1}(S\Lambda) = \Lambda$$

Let us consider two operators, A and B, which are both diagonalizable. If they commute, *they will share a common set of eigenvectors that represent a linearly independent basis in the vector space.* This is a powerful result. Should two operators be identified that have eigenvectors and commute, we know that they possess a common set of eigenvectors that simultaneously diagonalize both operators.

To see this, let us consider:

$$AB = BA$$

so that

$$[A, B] = 0.$$

Let us assume that A and B have a matrix, S, that simultaneously diagnolizes both. Keeping in mind that diagonal matrices, even defined over the field of complex numbers, always commute:

$$\begin{aligned} AB &= S\Lambda_1 S^{-1} S\Lambda_2 S^{-1} = S\Lambda_1\Lambda_2 S^{-1} \\ BA &= S\Lambda_2 S^{-1} S\Lambda_1 S^{-1} = S\Lambda_2\Lambda_1 S^{-1} \end{aligned}$$

Since diagonal matrices always commute, we see that $\Lambda_2\Lambda_1 = \Lambda_1\Lambda_2$ and thus $AB = BA$.

It is matrices that DON'T commute that suffer from Heisenberg's Uncertainty Principle, as we will see in the close of this section.

3.3.6 The Expectation Value and the Variance

The *expectation value* of an operator, A, applied to a vector, $|V\rangle$, is a number that represents the most likely value one would obtain by measuring the quantity represented by the operator. It is written as:

$$\langle A \rangle = \langle V | \, A \, | V \rangle \, .$$

For a Hermitian operator, this is a real-valued number. Since it's a number, when it appears in a calculation it can be moved anywhere in the calculation at no penalty (e.g. it doesn't have to be moved carefully, like a matrix).

The *variance* is the average of the typical variations of the measured values (obtained by applying A to the vector) from the mean (the expectation value). It is defined as:

$$\begin{aligned}(\Delta A)^2 &= \langle V|(A-\langle A\rangle)^\dagger(A-\langle A\rangle)|V\rangle \\ &= \langle V|\,(A-\langle A\rangle)^2\,|V\rangle \\ &= \langle (A-\langle A\rangle)^2\rangle.\end{aligned}$$

It is often convenient to define:

$$|X\rangle = (A-\langle A\rangle)\,|V\rangle$$

so that

$$(\Delta A)^2 = \langle X|X\rangle$$

Then we can similarly say:

$$(\Delta B)^2 = \langle Y|Y\rangle$$

and

$$\langle (B-\langle B\rangle)^\dagger(A-\langle A\rangle)\rangle = \langle Y|X\rangle.$$

We almost immediately can combine things we've learned above (the Schwartz Inequality) and come to an interesting formula, which we will use later:

$$\begin{aligned}(\Delta A)^2(\Delta B)^2 &= \langle X|X\rangle\langle Y|Y\rangle &(1)\\ &= |X|^2|Y|^2 &(2)\\ &\geq |\langle X|Y\rangle|^2. &(3)\end{aligned}$$

It is convenient to rewrite this inequality in a more suggestive form, starting from:

$$\begin{aligned}|\langle X|Y\rangle|^2 &= zz^* \geq \left(\frac{\langle X|Y\rangle - \langle Y|X\rangle}{2i}\right)^2 \\ &\geq \left(\frac{\langle (A-\langle A\rangle)(B-\langle B\rangle)\rangle - \langle (B-\langle B\rangle)(A-\langle A\rangle)\rangle}{2i}\right)^2 \\ &= \left(\frac{\langle AB\rangle - \langle A\rangle\langle B\rangle - \langle B\rangle\langle A\rangle - \langle A\rangle\langle B\rangle - \langle BA\rangle + A\langle B\rangle + B\langle A\rangle - \langle A\rangle\langle B\rangle}{2i}\right)^2 \\ &= \left(\frac{\langle AB\rangle - \langle BA\rangle}{2i}\right)^2 \\ &= \left(\frac{\langle [A,B]\rangle}{2i}\right)^2\end{aligned}$$

This finally lets us write:

$$\sqrt{(\Delta A)^2(\Delta B)^2} = \Delta A\Delta B \geq \frac{\langle [A,B]\rangle}{2i}.$$

There are far better ways to generalize this, but it will serve our purposes for this lecture.

3.4 A comment on function spaces

The concept of a vector space is not limited only to mathematical objects in a matrix representation. Functions of a continuous variable, like $f(x)$, can ALSO be part of a space with addition, scalar multiplication, etc. This kind of space is a Hilbert Space. There are relationships between the matrix representation and the function representation of a space; the Dirac Notation accommodates both of them.

- In the matrix representation, a ket is a column matrix; in a Hilbert Space, a ket is a function, like $f(x)$.
- In the matrix representation, an operator is a square matrix; in a Hilbert Space, an operator is a function that modifies $f(x)$ and returns another function in the Hilbert Space, $g(x)$.
- In the matrix representation, the inner product is a matrix multiplication that results in a real number; in a Hilbert Space, the inner product is an infinite sum over the products of two functions, $f^*(x)g(x)$, multiplied by a scale factor Δ (related to the step size of the sum), in the limit that $\Delta \to 0$. Thus:

$$\langle f|g\rangle = \int_{-\infty}^{\infty} f^*(x)g(x)\,dx.$$

- In the matrix representation, the expectation value is a matrix multiplication of a row vector, a square matrix, and a column vector, yielding a real number. In a Hilbert Space:

$$\langle f|A|g\rangle = \int_{-\infty}^{\infty} f^*(x)Ag(x)\,dx.$$

Either of these approaches is part of a strategy to utilizing the Schroedinger Wave Equation, and they will often be mixed as we move forward.

4 The Schroedinger Wave Equation (Non-Relativistic Quantum Mechanics)

It is impossible, in the context of this lecture series, to review in depth the Schroedinger Wave Equation. You are expected to have already seen this at least once in an undergraduate course. The fundamental equation of quantum mechanics is a re-statement of energy conservation using the Hamiltonian Formalism. That is, according to this formalism:

$$H = T + V = E$$

where H, the "Hamiltonian" of the system which completely describes energy in the system, is the sum of kinetic energy, T, and potential energy, V . This should be equal to the total energy of the system.

Consider a single-particle system. In classical mechanics, the kinetic energy of a particle can be written:

$$T = \frac{p^2}{2m}.$$

(I leave it as an exercise to the student to demonstrate this, starting from the earliest form we learn, $T = KE = \frac{1}{2}mv^2$). In quantum mechanics, momentum, P, is an *operator* whose action on the wave function of the single-particle system is to measure the momentum of the system, e.g.:

$$P\left|\psi(x,y,z)\right\rangle = p\left|\psi(x,y,z)\right\rangle.$$

In terms of linear algebra/matrix mechanics, we recognize this statement as an *eigenvalue equation*. We can write the operator, P, as a function in a Hilbert Space - a vector space for functions. The functional form of the momentum operator is

$$P = -i\hbar\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$$

and so we arrive at the operator form of the Schroedinger Wave Equation:

$$\begin{aligned} H\left|\psi\right\rangle &= E\left|\psi\right\rangle \\ (T+V)\left|\psi\right\rangle &= E\left|\psi\right\rangle \\ \left(\frac{1}{2m}P^2 + V\right)\left|\psi\right\rangle &= E\left|\psi\right\rangle \\ \left(-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}, \frac{\partial^2}{\partial y^2}, \frac{\partial^2}{\partial z^2}\right) + V\right)\left|\psi\right\rangle &= E\left|\psi\right\rangle \end{aligned}$$

The devil is in the details, of course. Usually, once you learn this, you then make your pencil bleed by having to solve for the energy eigenvalues given a potential (e.g. the harmonic oscillator potential, a square well, etc.) and a wave function. We will not do this, but instead skip on into orbital angular momentum (e.g. consider system like a single electron trapped in the electric potential of a central nucleus, which can execute orbital motion in the presence of that potential).

In-Class Exercise: The Momentum of a Plane Wave

Consider a plane wave, whose wave function is given by:

$$\left|\psi(x,y,z,t)\right\rangle = Ae^{-i\hbar(kx-\omega t)}.$$

What is the momentum of this plane wave?

5 Review of Angular Momentum

In classical physics, angular momentum is defined as the cross-product of the lever arm direction, $\vec{r}$, at a point (with respect to the center of rotational motion) with the linear momentum, $\vec{p}$, at that same point:

$$\vec{L} = \vec{r} \times \vec{p}.$$

We know the form of the momentum:

$$P\left|\psi(x,y,z)\right\rangle = -i\hbar\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)\left|\psi(x,y,z)\right\rangle.$$

There are relationships between the components of the angular momentum. If we write out the cross-products that yield each component:

$$\begin{aligned} L_x &= r_y p_z - r_z p_y \\ L_y &= r_z p_x - r_z p_z \\ L_z &= r_x p_y - r_y p_x \end{aligned}$$

we can then ask and answer questions like:

- Does the angular momentum operator commute with the Hamiltonian? What does this imply?
- Since, in quantum mechanics, momentum is an operator and not simply a number, do the components of angular momentum *commute* with one another? That is, can we say that $L_x L_y = L_y L_x$ and the result of any measurement of the x-component of angular momentum, then the y-component of angular momentum, is the same if we measure first y and then x? If not, what does this imply?

In-Class Exercise: Commutation Relations of the Angular Momentum Operators

Demonstrate that the commutation relations of the components of the angular momentum operators are, in fact:

$$[L_i, L_j] = i\hbar L_k.$$

In other words, show that the components DO NOT commute and the order of operation (measure x, then y) affects the outcome of the measurement. From this, we can derive the following useful relationship, which we take as given for the remainder of this lecture series:

$$\vec{L} \times \vec{L} = i\hbar\vec{L}.$$

5.1 The Commutation of the Hamiltonian and the Angular Momentum Operator

It can be shown that, if the potential part of the Hamiltonian consists of a *spherically symmetric potential*, then the Hamiltonian commutes with the Angular Momentum Operator; that is,

$$[L^2, H] = 0.$$

and also that

$$[L_i, H] = 0$$

What does this imply?

As discussed earlier in the "Basic Mathematics" section, commutation of two operators implies that they each possess a common set of eigenvectors that simultaneously diagonalize both operators (yielding common eigenvalues). We can describe these solutions using by eigenvectors (wave functions) of the form $|\psi(E,\alpha)\rangle$, where E and α here denote the eigenvalues of H and L^2, respectively.

5.2 The Non-Commutation of Angular Momentum Components and Information

The fact that we cannot, in quantum mechanics, commute all of the individual components of angular momentum has consequences.

- We can only specify with certainty one of the three components at any one time. It is conventional to choose L_z.
- This is, in a deep sense, directly related to the Heisenberg Uncertainty Principle. Having exactly specified L_z, we can ask if there is an inequality relationship between the remaining two components of the angular momentum, ala

$$\Delta L_x \Delta L_y \geq \frac{\langle [L_x, L_y] \rangle}{2i}.$$

(in other words, is the right-hand side non-zero, so that this is more than a trivial inequality?). To answer this, we grind through the calculation:

$$\begin{aligned} \langle [L_x, L_y] \rangle &= \langle \psi | L_x L_y - L_y L_x | \psi \rangle \\ &= \langle \psi | (i\hbar L_z) | \psi \rangle \\ &= i\hbar \langle \psi | L_z | \psi \rangle \\ &= i\hbar \langle L_z \rangle. \end{aligned}$$

So we arrive at the statement of the Heisenberg Uncertainty Principle for angular momentum:

$$\Delta L_x \Delta L_y \geq \frac{\hbar}{2} \langle L_z \rangle \geq \frac{\hbar}{2}.$$

We conclude from this that if one specifies the angular momentum along the z-direction, one has NO control in specifying the angular momentum along both the x- and y-components with absolute precision.

5.3 The Eigenvalues of L_z

We can write the $L_z$operator in the space of cartesian coordinates (x,y,z) as:

$$L_z = -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right).$$

It is more convenient, however, to express this in spherical coordinates - especially because in quantum mechanics we are often interested in spherically symmetric potentials, and these problems simplify in spherical coordinates.

In-Class Exercise: Transformation of Coordinate System
Transform the Cartesian-space form of L_z to the spherical-coordinate form. That is, express L_z not in (x,y,z) but in (ρ,θ,ϕ).

The correct expression in spherical coordinates is quite simple:

$$L_z = -i\hbar \frac{\partial}{\partial \phi}.$$

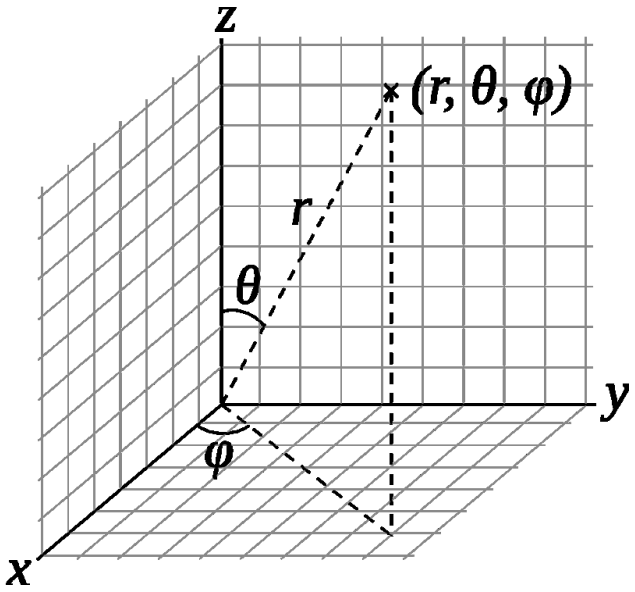


Figure 4: The relationship between Cartesian coordinates and spherical coordinates. Figure is from Ref. [22].

We can now setup and solve the eigenvalue problem, using the functional form of the operator:

$$L_z \left|\psi(\rho,\phi)\right\rangle = l_z \left|\psi(\rho,\phi)\right\rangle \longrightarrow -i\hbar\frac{\partial}{\partial\phi}\psi(\rho,\phi) = l_z\psi(\rho,\phi).$$

This is a first-order differential equation, and by inspection (that is, with some experience in solving these under your belt!), you can simply write the solution as:

$$\psi(\rho,\phi) = R(\rho)e^{il_z\phi/\hbar}.$$

At first glance, it appears that l_z can be any complex number. We must impose the Hermiticity requirement on the problem in order to come to a complete understanding of the eigenvalues:

$$\langle\psi_1| L_z |\psi_2\rangle = \langle\psi_2| L_z |\psi_1\rangle^* .$$

If we write this in the coordinate basis:

$$\int_0^\infty \int_0^{2\pi} \psi_1^* \left(-i\hbar\frac{\partial}{\partial\phi}\right)\psi_2 \; \rho\, d\rho\, d\phi = \left[\int_0^\infty \int_0^{2\pi} \psi_2^* \left(-i\hbar\frac{\partial}{\partial\phi}\right)\psi_1 \; \rho\, d\rho\, d\phi\right]^* . \tag{4}$$

To solve this, integration by parts is required:

$$\int_a^b u\, dv = uv\Big|_a^b - \int_a^b v\, du$$

where in our case we identify:

$$\begin{aligned} u &= \psi_2^* \\ dv &= -i\hbar\frac{\partial}{\partial\phi}\psi_1 d\phi = -i\hbar d\psi_1 \\ du &= d\psi_2^* = \frac{\partial}{\partial\phi}\psi_2^* d\phi \\ v &= -i\hbar\psi_1 \end{aligned}$$

We can then write the right-hand side of the original equality as:

$$\begin{aligned}\left[\int_0^\infty \int_0^{2\pi} \psi_2 \left(-i\hbar\frac{\partial}{\partial\phi}\right)\psi_1 \, \rho \, d\rho \, d\phi\right]^* &= \left[\int_0^\infty \rho \, d\rho \, (-i\hbar\psi_2^*\psi_1)|_0^{2\pi} - \int_0^\infty \int_0^{2\pi} \rho \, d\rho \left(-i\hbar\psi_1\frac{\partial}{\partial\phi}\psi_2^* d\phi\right)\right]^* \\ &= \int_0^\infty \rho \, d\rho \, (-i\hbar\psi_2\psi_1^*)|_0^{2\pi} + L.H.S.\end{aligned}$$

where "L.H.S." is the left-hand side of Equation 4. The original L.H.S. and this term cancel each other, leaving:

$$\int_0^\infty \rho \, d\rho \, (-i\hbar\psi_2\psi_1^*)|_0^{2\pi} = 0$$

We can simply then write a relationship between the two integrands, since every other part of the integral is identical, that satisfies this above requirement:

$$\psi_2(\rho, 2\pi)\psi_1^*(\rho, 2\pi) - \psi_2(\rho, 0)\psi_1^*(\rho, 0) = 0.$$

The only way to satisfy this equation is if $\psi(\rho, 2\pi) = \psi(\rho, 0)$. This imposes a boundary condition on the angular part of the wave function:

$$1 = e^{2\pi i l_z/\hbar}.$$

We thus find that

1. l_z must be a real number
2. l_z can only take integer values if it is to satisfy this equation

Thus:

$$l_z = m\hbar, \; m = 0, \pm 1, \pm 2, ...$$

and we learn that the z-component of angular momentum is predicted to be a quantized quantity. We can label m to be the *magnetic quantum number*, the true quantum number of the state describing the projection of orbital angular momentum along the z-direction.

5.4 The Eigenvalues of L^2

It can be shown that there is another operator, independent of L_z, that not only also commutes with the Hamiltonian (for azimuthally symmetric wave functions) but also with L_z itself. That operator is the square of the total angular momentum vector, $L^2 = L_x^2 + L_y^2 + L_z^2$. Once we specify the eigenvalues of this operator, we can completely specify the angular momentum state of a system.

What are the eigenvalues of this operator? We can begin by writing:

$$L^2 \, |\psi(\alpha, m)\rangle = \alpha \, |\psi(\alpha, m)\rangle$$

where m are the eigenvalues of the L_z operator and α are the eigenvalues of the L^2 operator. I am intentionally being a bit careful about not assuming that the eigenvalues are just ℓ^2 on the right-hand side of this equation; as we will see, applying the momentum operator twice in succession doesn't merely yield the square of a single number.

For the next step, it will be convenient to rewrite L_x and L_y in terms of two other operators,

$$\begin{aligned}L_+ &\equiv L_x + iL_y \\ L_- &\equiv L_x - iL_y\end{aligned}$$

and then ask the question: what are the commutation relations of these operators with L_z? To answer this:

$$\begin{aligned}[L_z, L_+] &= [L_z, L_x] + i[L_z, L_y] \\ &= i\hbar L_y + i\left(-i\hbar L_x\right) \\ &= \hbar\left(L_x + iL_y\right) = \hbar L_+\end{aligned}$$

Likewise,

$$\begin{aligned}[L_z, L_-] &= [L_z, L_x] - i[L_z, L_y] \\ &= i\hbar L_y - i\left(-i\hbar L_x\right) \\ &= -\hbar\left(L_x - iL_y\right) = -\hbar L_-\end{aligned}$$

We can then consider what it means to have these new operators, L_+ and L_-, act on a definite state of $|\psi(m)\rangle$. Specifically, we can ask, "What will L_z measure after $L_\pm$ acts on a definite state of $|\alpha, \beta\rangle$?" Let's give this a try:

$$\begin{aligned}L_z\left(L_+ |\psi(m)\rangle\right) &= \left([L_z, L_+] + L_+ L_z\right)|\alpha, \beta\rangle \\ &= \left(\hbar L_+ + L_+ L_z\right)|\alpha, \beta\rangle \\ &= \hbar L_+ |\alpha, \beta\rangle + L_+ \beta |\alpha, \beta\rangle \\ &= (\beta + \hbar)\left(L_+ |\alpha, \beta\rangle\right)\end{aligned}$$

From this, we conclude that:

$$L_z\left(L_+ |\alpha, \beta\rangle\right) = L_z |\alpha, \beta + \hbar\rangle$$

That is, the act of applying $L_+$to a state of definite m, $|\alpha, \beta\rangle$, and then measuring the z-projection of its total angular momentum is EQUIVALENT to having just applied the L_z operator to a state $|\alpha, \beta + \hbar\rangle$. We can then see that L_+ is the "Raising Operator" of angular momentum projection along the z-axis. Similarly, an equivalent exercise reveals that L_- is the Lowering Operator of angular momentum, declining the projection along the z-axis by one unit. For a fixed amount of total orbital angular momentum, it is possible to move m through all its possible values up to a maximum value and down to a minimum value. If the maximum value is β_{max}, the minimum value must be $-\beta_{max}$ and the distance between the minimum and maximum values of m is given by $2\beta_{max}$. There is a VCR joke in here someplace, but I'm probably the only one old enough to get it.

We can also determine the commutation relation between the raising and the lowering operator:

$$\begin{aligned}[L_+, L_-] &= L_+ L_- - L_- L_+ \\ &= (L_x^2 - iL_x L_y + iL_y L_x + L_y^2) - (L_x^2 + iL_x L_y - iL_y L_x + L_y^2) \\ &= -2iL_x L_y + 2iL_y L_x \\ &= -2i(L_x L_y - L_y L_x) \\ &= -2i(i\hbar L_z) \\ [L_+, L_-] &= 2\hbar L_z\end{aligned}$$

You can also demonstrate that $[L^2, L_+] = 0$ and $[L^2, L_-] = 0$. From this, you can see that:

$$L^2 L_+ |\alpha, \beta\rangle = L_+ L^2 |\alpha, \beta\rangle = \alpha L_+ |\alpha, \beta\rangle .$$

We should now ponder the two equations involving raising/lowering operators and L^2 and L_z,

$$\begin{aligned}L^2 L_\pm |\alpha, \beta\rangle &= \alpha L_\pm |\alpha, \beta\rangle \\ L_z L_\pm |\alpha, \beta\rangle &= (\beta + \hbar) L_\pm |\alpha, \beta\rangle .\end{aligned}$$

We see that if we were to ask, what happens when

$$L_+ |\alpha, \beta\rangle$$

we would be able to conclude that $L_+ |\alpha, \beta\rangle \propto |\alpha, \beta + 1\rangle$, or more mathematically:

$$\begin{aligned} L_+ |\alpha, \beta\rangle &= C(\alpha, \beta)_+ |\alpha, \beta + 1\rangle \\ L_- |\alpha, \beta\rangle &= C(\alpha, \beta)_- |\alpha, \beta - 1\rangle \end{aligned}$$

To solve the original problem in which we were interested - the eigenvalues of L^2 - it is convenient to write:

$$\begin{aligned} L^2 &= L_z^2 + L_x^2 + L_y^2 \\ &= L_z^2 + \frac{1}{2}(L_+ L_- + L_- L_+). \end{aligned}$$

We can also do a bit more mathematical work and see that:

$$\begin{aligned} L^2 &= L_z^2 + \frac{1}{2}(L_+ L_- - L_- L_+ + L_- L_+ + L_- L_+) \\ &= L_z^2 + \frac{1}{2}([L_+, L_-] + 2L_- L_+) \\ L^2 - L_z^2 &= \frac{1}{2} 2\hbar L_z + L_- L_+ \\ L^2 - L_z^2 - \hbar L_z &= L_- L_+ \end{aligned}$$

With all of these relationships in mind, let's move to the real problem we want to solve.

Let us apply the operator $L_- L_+$ to a definite state of $|\alpha, \beta\rangle = |\alpha, \beta_{max}\rangle$:

$$L_- L_+ |\alpha, \beta_{max}\rangle = 0.$$

Why? The ladder operator cannot raise β above its maximum value. But, then, it must also be true that:

$$\begin{aligned} (L^2 - L_z^2 - \hbar L_z) |\alpha, \beta_{max}\rangle &= 0 \\ \alpha - \beta_{max}^2 - \hbar \beta_{max} &= 0 \\ \alpha &= \beta_{max}^2 + \hbar \beta_{max} \end{aligned}$$

So we have our first glimpse of an eigenvalue of L^2, and we see that it's related to the maximum value that β can take. What about considering a state where the minimum possible value of β is present? In that case, it must be true that:

$$\begin{aligned} L_- |\alpha, \beta_{min}\rangle &= 0 \\ L_+ L_- |\alpha, \beta_{min}\rangle &= 0 \\ (L^2 - L_z^2 + \hbar L_z) |\alpha, \beta_{min}\rangle &= 0 \\ \alpha - \beta_{min}^2 + \hbar \beta_{min} &= 0 \\ \alpha &= \beta_{min}(\beta_{min} - \hbar) \end{aligned}$$

And so it must also be true that $-\beta_{min} = \beta_{max}$, by these two relations. With all of this information, we can now solve for the eigenvalues of L^2.

We have learned the following:

1. The eigenvalues of L^2 are related to the integer eigenvalues of L_z, $\ell_z = m\hbar$.

2. While L_z can have a range of eigenvalues, $m\hbar = (0, \pm 1, \pm 2, ...)\hbar$, the eigenvalues of L^2 are fixed for a given set of eigenvalues for L_z by the relationship $\alpha = \beta_{max}(\beta_{max} + \hbar)$.

It is convention to say that a system is prepared in a state whose *orbital angular momentum quantum number*, ℓ, is related to the z-*projection of the total angular momentum,* m, by $\ell = m_{max}$. The actual total orbital angular momentum of the system is then given by:

$$\sqrt{\alpha} = \sqrt{\ell\hbar(\ell\hbar + \hbar)} = \sqrt{\ell(\ell+1)}\hbar.$$

Note that this is NOT required to be an integer; the only quantity that is guaranteed to be integer-quantized is the z-component of the total orbital angular momentum. We can then say that for a given ℓ, we have a range of allowed projections of L_z:

ℓ	m
0	0
1	$0, \pm 1$
2	$0, \pm 1, \pm 2$

etc.

5.5 Solutions to the Hydrogen Atom

It is instructive to look at the solutions to the hydrogen atom (a full, 3-D spherical potential), to get a sense of what the angular momentum wave functions will look like for two particles in a bound state (a common problem in nature).

$$\psi(\rho, \theta, \phi) = \sqrt{\left(\frac{2}{na_0}\right)^3 \left(\frac{(n-\ell-1)!}{2n(n+\ell)!}\right)} e^{-\rho/2} \rho^\ell L^{2\ell+1}_{n-\ell-1}(\rho) Y_l^m(\theta, \phi)$$

where

$$L^{2\ell+1}_{n-\ell-1}(\rho) = \frac{\rho^{-(2\ell+1)} e^\rho}{(n-\ell-1)!} \frac{d^{n-\ell-1}}{dx^{n-\ell-1}} \left(e^{-\rho} \rho^{n+\ell}\right)$$

is an Laguerre polynomial, describing the radial structure of the orbit with quantum numbers (n, ℓ, m) and

$$Y_\ell^m(\theta, \phi) = \sqrt{\frac{(2\ell+1)}{4\pi} \frac{(\ell-m)!}{(\ell+m)!}} P_\ell^m(\cos\theta) e^{im\phi}$$

are the normalized spherical harmonics containing the associated Legendre Polynomials,

$$P_\ell^m(\cos\theta) = (-1)^m (1 - \cos^2\theta)^{m/2} \frac{d^m}{d\cos\theta^m} \left(\frac{1}{2^\ell \ell!} \frac{d^\ell}{d\cos\theta^\ell} \left[\left(\cos^2\theta - 1\right)^\ell\right]\right).$$

Some of the spherical harmonics are visualized in Fig 5.

We see that the angular momentum controls the structure of the orbits of an electron, giving them shape. This seems like a trivial conclusion - of course, angular momentum should be the thing that controls the shapes since it has to do with the orbits themselves - but this basic idea will carry forward into things like particle decay. When a particle decays, the parent contains a certain amount of angular momentum (potentially both in the form of orbital and spin angular momentum), and any orbital momentum present in the final state control the angular structure of the outgoing particles. How that orbital angular momentum manifests and is partitioned depends on the particular details of the decay, and the spin angular momentum of the final-state particles.

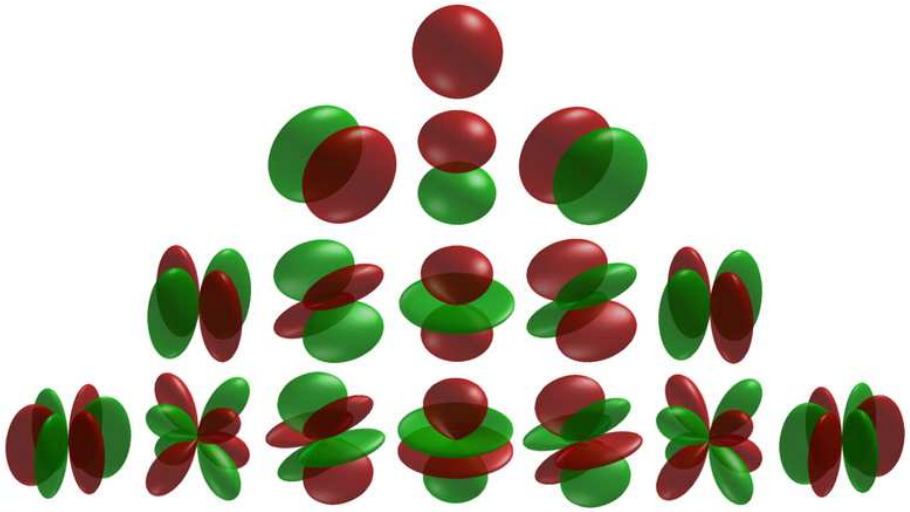


Figure 5: Visualization of some of the spherical harmonics. From top to bottom, the rows represent $\ell = 0, 1, 2, 3$, while the entries from left-to-right in each row represent $m = -\ell, -\ell+1, ..., 0, ..., \ell-1, \ell$. From Ref. [23].

6 Spin Angular Momentum

6.1 Describing spin using vector spaces

The observation of spectral lines in the alkali metals reveal that for each principle quantum number, n, there are two corresponding, finely spaced spectral lines. This implies a two-state phenomenon as regard the electron, further splitting each spectral line into two distinct lines. We'll see this in the math later. For now, let us assume we need to describe a two-state angular momentum phenomenon. Denote the wave function of the electron as:

$$|\psi\rangle$$

and for now let us assume that the portion of the wave function associated with this additional set of states is independent of the other parts of the wave function (e.g. orbital angular momentum), so that

$$|\psi\rangle = |\psi(\pm, \ell, m)\rangle\,.$$

Here, I have denoted the new two-state portion of the wave function as $|\psi(\pm)\rangle$. We can short-hand this as $|\pm\rangle$. We can now explicitly write the two independent states as orthonormal basis vectors in a two-component space:

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

These two-component objects are not vectors, and they are not scalars; they are called "spinors." One then then construct an arbitrary state (spinor) from these basis spinors:

$$|\chi\rangle = a\,|+\rangle + b\,|-\rangle\,.$$

Let us now explore the features of this new quantity, which is called "spin angular momentum" or just "spin" for short. Please note: while the original historical models of spin actually did include a spinning electron, there is no actual mechanical motion of the electron present that results in this property. *Spin* is an inherent quantum property which we can describe using the mechanical analog, but for which there is *no actual mechanical equivalent* in classical physics. You had a homework problem on that helped illustrate the point.

It is an empirical fact that if one prepares an atom in a state of zero total orbital angular momentum and one then measures the angular momentum of the electron, it still presents two non-zero values of its angular momentum component along the z-direction:

$$S_z = \pm\frac{1}{2}\hbar.$$

Based on this observation, we can then determine the form of the operator required to measure the z-component of the spin angular momentum of the electron. We do this in two steps.

1. Prepare a state of pure "spin-up," $s_z = \frac{1}{2}\hbar$, and measure that component using the z-component of the spin operator:

$$S_z\,|\chi\rangle = +\frac{1}{2}\hbar\,|\chi\rangle \longrightarrow \begin{pmatrix} s_1 & s_2 \\ s_3 & s_4 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = +\frac{1}{2}\hbar\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

   If we solve this, we find that $s_1 = +\frac{\hbar}{2}$ and $s_3 = 0$

2. Prepare a second state of pure "spin-down," $s_z = -\frac{\hbar}{2}$. Measure that component:

$$S_z\,|\chi\rangle = -\frac{1}{2}\hbar\,|\chi\rangle \longrightarrow \begin{pmatrix} s_1 & s_2 \\ s_3 & s_4 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = +\frac{1}{2}\hbar\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

   Again, if we solve this for the unknown components of the spin projection operator along the z-direction, we find $s_3 = 0$ and $s_4 = -\frac{\hbar}{2}$.

Assembling all of the pieces, we find the form of the S_z operator:

$$S_z = \frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

It can be shown that the operators for projecting the spin of a particle along the x- and y-directions are given by:

$$S_x = \frac{1}{2}\hbar \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; S_y = \frac{1}{2}\hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

We can write generally that

$$\vec{S} = \frac{\hbar}{2}\vec{\sigma}.$$

Please note that the designation of which direction is x,y,z is conventional; we could, historically, have talked instead about measuring the spin projection of an electron along the x-direction and defined that as the direction of the magnetic field gradient in, say, a Stern-Gerlach-type experiment. But we adopt this convention (where the z-axis is, in fact, that direction) and proceed consistently.

6.2 Total Angular Momentum

Since orbital angular momentum alone is insufficient to describe the total angular momentum of, say, an atom, we must reformulate the operator language we developed purely for orbital angular momentum. In doing so, we can describe total angular momentum in a system of particles. Let us define:

$$\vec{J} = \vec{L} + \vec{S},$$

where $\vec{J} = (J_x(= L_x + S_x),\ J_y(= L_y + S_y),\ J_z(= L_z + S_z))$ and each of the elements of this operator is also an operator. We can quickly reuse our old definitions for commutation relations of orbital angular momentum, ladder operators, etc.

$$\begin{aligned}
J_z \left|j,m\right\rangle &= m\hbar \left|j,m\right\rangle \\
J^2 \left|j,m\right\rangle &= j(j+1)\hbar^2 \left|j,m\right\rangle \\
J_\pm \left|j,m\right\rangle &= \sqrt{(j \mp m)(j \pm m + 1)}\hbar \left|j, m \pm 1\right\rangle \\
[J^2, J_z] &= 0 \\
[J_+, J_-] &= 2\hbar J_z \\
[J_\pm, J_z] &= \pm\hbar J_\pm
\end{aligned}$$

We can also quickly recognize, by thinking about total angular momentum as the vector sum or orbital and spin angular momenta, that:

$$\begin{aligned}
j &= \ell + s \\
m &= m_\ell + m_s.
\end{aligned}$$

We see then that $j = 0, \pm\frac{1}{2}, \pm 1, \pm\frac{3}{2}, \pm 2, ...$ and the total magnetic quantum number is given by the range: $m = -j,\ -j+\frac{1}{2},\ -j+1,\ ...,\ 0,\ ...,\ j-1,\ j-\frac{1}{2},\ j$.

We also need to define a new operator, which results from:

$$J^2 = L^2 + 2\vec{L}\cdot\vec{S} + S^2.$$

We have a "spin-orbit term" present in the total angular momentum squared:

$$\begin{aligned}
\vec{L}\cdot\vec{S} &= \frac{1}{2}\left(J^2 - L^2 - S^2\right) \\
\vec{L}\cdot\vec{S}\left|\ell, m_\ell, s, m_s\right\rangle &= \frac{\hbar^2}{2}\left[j(j+1) - \ell(\ell+1) - s(s+1)\right]\left|\ell, m_\ell, s, m_s\right\rangle.
\end{aligned}$$

6.3 Manipulating Angular Momentum

Consider a current loop in a classical physics setting. If exposed to an external magnetic field, $\vec{B}$, a rotational force is exerted on the loop. This torque is given by:

$$\vec{\tau} = \vec{\mu} \times \vec{B}$$

where $\vec{\mu}$ is the *magnetic dipole moment* of the loop. In the Bohr or Sommerfeld models of the atom, the electron in orbit around the central nucleus is analogous to a classical current loop. We might then start from the classical effect of a magnetic dipole interacting with an external magnetic field, and then extend that into the quantum realms of orbital angular momentum and spin angular momentum.

6.3.1 Classical Model

Consider a square loop of current, I. The magnetic moment is given by:

$$\vec{\mu} \quad = \quad \frac{I \cdot A}{c} \hat{n}$$

where A is the area of the loop, c is the speed of light, and $\hat{n}$ is a unit vector that is determined by a right-hand rule: curl your fingers on your right hand in the direction of current flow (the direction positive charge is flowing in the circuit) and your thumb then points in the direction $\hat{n}$.

Such a current loop, subjected to an external magnetic field, experiences the torque given above; this torque tends to rotate the loop until the magnetic moment is parallel to the external magnetic field, and then the rotation stops. The interaction energy is given by:

$$\mathcal{H}_{int} = \int T(\theta)\, d\theta = \int \mu B \sin(\theta)\, d\theta = -\mu B \cos\theta = -\vec{\mu} \cdot \vec{B}.$$

While conceived using a square loop of current, the formulas apply just as well to circular flows of current. So, let's consider a very simple model of an atom.

Imagine a single electron, with charge e and mass m, orbiting in a circle under the influence of a central Coulomb potential (e.g. due to a single proton). The current associated with the charge is:

$$I = \frac{\Delta Q}{\Delta t} = \frac{q}{(2\pi r)/v} = \frac{qv}{2\pi r}$$

where v is the speed of the electron and r is the orbital radius. The magnetic moment of this classical, "toy" atom is then:

$$\mu = \frac{IA}{c} = \frac{qv}{2\pi r}\frac{\pi r^2}{c} = \left(\frac{q}{2mc}\right) mvr = \frac{q}{2mc} L$$

where L is the angular momentum, $L = I\omega = mr^2 \frac{v}{r} = mvr$, of a single orbiting mass. Thus the magnetic moment is related to the single-particle orbital angular momentum, and we have this prefactor which is written as

$$\gamma = \frac{q}{2mc}$$

and is known as the "gyromagnetic ratio" of the electron. In general $\gamma = \frac{\mu}{L}$, and here in this toy model we have solved exactly for γ.

6.3.2 Quantum Model

In non-relativistic quantum mechanics, one writes the Hamiltonian for a charged particle under the influence of an external magnetic field in terms of the momentum operator and the vector potential, where:

$$\vec{B} = \vec{\nabla} \times \vec{A}$$

The Hamiltonian is then written as:

$$H = \frac{1}{2m}(\vec{P} - q\vec{A})^2$$

If one solves this for the interaction part of the Hamiltonian, which are the cross-terms involving $\vec{A} \cdot \vec{P}$, one finds that the classical result holds:

$$\vec{\mu} = \frac{q}{2Mc}\vec{L}.$$

where M is the mass of the particle (to avoid confusing it with m, the quantum number of the z-projection of angular momentum).

We know that angular momentum is quantized along the z-direction, so that if we then consider the z-component of the magnetic moment we find:

$$\mu_z = \frac{q}{2Mc} m\hbar \ (m = 0, \ \pm 1, ...) \, .$$

The quantity $q\hbar/2Mc$ is referred to as the *Bohr Magneton*, and for an electron is found to be:

$$\frac{q\hbar}{2Mc} \approx 0.6 \times 10^{-8} \text{eV/G}.$$

6.3.3 Spin Magnetic Moment

There is no classical analog for spin - it is merely intrinsic angular momentum specific to a particle. However, we can assume that, since in all other ways it manifests as angular momentum, it too must have a relationship to a magnetic moment:

$$\vec{\mu} = \gamma \vec{S}.$$

We can then write γ in a form reminiscent of the gyromagnetic ratio for the case of orbital angular momentum:

$$\gamma = g\frac{q}{2Mc}$$

where g is a factor that needs to be determined experimentally - it is not predicted by this framework.

The interaction term in the Hamiltonian between an external magnetic field and this intrinsic magnetic moment is then:

$$\mathcal{H}_{int}^{spin-B} = -\vec{\mu} \cdot \vec{B} = \frac{ge}{2Mc}\vec{S} \cdot \vec{B} = \frac{ge\hbar}{4Mc}\vec{\sigma} \cdot \vec{B}.$$

We see that the intrinsic magnetic moment due to spin is just $g/2$ Bohr Magnetons. Experimentally, $g \approx 2$. Making the approximation that $g = 2$, we observe that the intrinsic magnetic moment due to spin is TWICE that for orbital angular momentum. The fact that g is not exactly equal to 2 is important, and deeply connected to the more fundamental theory of nature - quantum field theory, and specifically quantum electrodynamics. It is possible, in that more fundamental model of nature, to calculate g from first principles. The current measurement of g and the theoretical calculation agree very well. The experimentally measured value is expressed in terms of its deviation from 2:

$$a = \frac{g-2}{2} = 0.0115965218073(28),$$

where the uncertainty is in the last two decimal places and is given in the parentheses. It is known to better than 1 part in 1 billion.

The measurement of the magnetic moment of particles, such as the electron, the muon, and the tau lepton, are not only tests of the Standard Model of Particle Physics but a means to probe for physics beyond

the Standard Model. For instance, additional particles, not described by the Standard Model but present in nature, can participate in self-interactions (higher-order Feynman diagrams) of the electron, muon, and tau lepton and influence the magnetic moment's value. The measurement of the muon magnetic moment is currently an area of hot pursuit, as the best measured value is not in perfect agreement with the calculated value. The predicted magnetic moment of the tau lepton has been calculated (c.f. Ref. [24]) but to date no direct measurement of it has been possible due to the very short lifetime of the tau lepton.

6.3.4 A Comment: Stern-Gerlach-style Experiments

A question for you right now would be the following:

- Why does the Stern-Gerlach experiment employ a magnetic field gradient rather than just a uniform magnetic field?

The answer is that a uniform field can only rotate the magnetic moments to align with the field. Since the magnetic moments of electrons have two possible orientations along the z-axis - along or against - half the electrons should align with and half against the field. But this would not split the beam, because once the magnetic moments are aligned properly with the field there is no remaining force on the atoms or electrons.

Instead, a magnetic field gradient is needed to then push the little internal magnets of the atoms/electrons in a specific direction in the field. This can be seen as follows:

$$\vec{F} = -\vec{\nabla}\mathcal{H}_{int} = \vec{\nabla}(\vec{\mu}\cdot\vec{B}) = (\vec{\mu}\cdot\vec{\nabla})\vec{B} = \mu_z\frac{\partial B_z}{\partial z}\hat{k}.$$

This equation tells us that a gradient is required to exert a force on the tiny magnetic moments of the atom/electron. It is this tiny force that allows for the splitting of the beam in a Stern-Gerlach-style experiment. One expects the splitting to occur along the magnetic field axis.

It is important to note that even if one prepares the atoms in a S-G-style experiment in a state $L = 0$ so that $L_z = 0\hbar$, it is still possible to split the beam if:

- The atom contains an odd number of electrons in its valence shell (its outermost shell). This leaves an unpairs spin angular momentum, and this unpaired spin angular momentum has no other spin to compensate for its orientation along/against the magnetic field gradient
- The electron must contain an internal unit of angular momentum, to allow for an internal magnetic moment to interact with the external magnetic field gradient.

Since the above occurs in an S-G-style experiment with L=0, it allows us to infer the presence of internal angular momentum and measures its value.

6.4 Addition of Spin Angular Momentum

We have so far been focused on very, very simple systems containing only a single particle with spin angular momentum, or spin and orbital angular momentum. But what happens for more realistic systems (e.g. where you have multiple particles in various states of spin and orbital angular momentum)? For instance, we know experimentally that quarks never appear by themselves in experiments; they always appear, when detectable in a final state, in bound pairs (mesons, like the pion) or triplets (baryons, like the proton or neutron). Thus the structure of matter (e.g. the proton, or pions that participate in strong interactions) fundamentally depends on two- and three-particle systems. Regarding spin angular momentum, how does this business work then?

Consider a simple extension of the one-particle system: a two-particle system. Let each particle have spin angular momentum $s_i = \frac{1}{2}\hbar$ and let each have possible z-axis projections of its spin angular momentum, $s_{zi} = \pm\frac{1}{2}\hbar$. We have two particles, each with their own spin eigenvectors; this is a two-particle Hilbert space, requiring four total vectors to span the space. We might naively write the four vectors thus:

$$|\psi_1, \psi_2\rangle = |s_1, m_1, s_2 m_2\rangle$$

so that the four vectors are:

$$\begin{aligned}
|1\rangle &= \left|\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}\right\rangle \equiv |{+}{+}\rangle \\
|2\rangle &= \left|\frac{1}{2},-\frac{1}{2},\frac{1}{2},\frac{1}{2}\right\rangle \equiv |{-}{+}\rangle \\
|3\rangle &= \left|\frac{1}{2},\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right\rangle \equiv |{+}{-}\rangle \\
|4\rangle &= \left|\frac{1}{2},-\frac{1}{2},\frac{1}{2},-\frac{1}{2}\right\rangle \equiv |{-}{-}\rangle
\end{aligned}$$

Let's see how far we can get with this hypothesis.

We will also need spin operators for this new space. It is straight-forward to show that:

$$\vec{S} = \vec{S}_1 + \vec{S}_2$$

yields exactly what we need, with all the necessary properties for our total spin operator. We then have:

$$S_z = S_{1z} + S_{2z}.$$

Let us then proceed to find the matrix elements of this operator. We have to consider operations like

$$S_z\,|{+}{+}\rangle = (S_{z1} + S_{z2})\,|{+}{+}\rangle = (\frac{1}{2} + \frac{1}{2})\hbar\,|{+}{+}\rangle = \hbar\,|{+}{+}\rangle\,.$$

We can see immediately that, since this operator leaves the ket unchanged all off-diagonal elements of S_z will vanish. But we also see that two of the on-diagonal elements will be idential, leading to a degeneracy in the space:

$$S_z = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \hbar.$$

We have two eigenvectors, $|{+}{-}\rangle$ and $|{-}{+}\rangle$, that yield the same eigenvalue. The basis vectors we have chosen naively clearly span the vector space - the matrix is diagonal. So maybe we have chosen, if accidentally, the correct basis vectors for our two-particle Hilbert space.

We must then turn to the S^2 operator. Let us compute the matrix elements of this operator; they are not immediately obvious (e.g. by inspection). We see that

$$S^2 = S_1^2 + S_2^2 + 2\vec{S}_1 \cdot \vec{S}_2.$$

which makes this a VERY complicated operator with cross-terms between the two particles. You can write this more simply as follows:

$$S^2 = S_1^2 + S_2^2 + 2\left(S_{1z}S_{2z} + \frac{1}{2}\left(S_{1+}S_{2-} + S_{1-}S_{2+}\right)\right).$$

One can then brute-force compute the matrix elements in our present basis, and in doing so we find:

$$S^2 = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \hbar^2.$$

This is very interesting. We see that the middle block of the matrix is not diagonal. What does this mean?

It means that we have *NOT IDENTIFIED THE COMMON BASIS OF THE TWO-PARTICLE HILBERT SPACE THAT SIMULTANEOUSLY DIAGONALIZES* S_z and S^2. That is why writing down the "naive" eigenvectors was a bit too naive. The goal in any multiparticle system is to identify those eigenvectors that do simultaneously diagonalize both angular momentum operators, allowing them to commute. We could have foreseen that this would be a problem slightly earlier, since the default expression for S^2, which includes a lot of cross-terms, does not automatically commute with S_z. If we can simultaneously diagonalize both matrices (by finding their common eigenbasis), we can guarantee that they commute.

We might try writing linear combinations of the kets $|+-\rangle$ and $|-+\rangle$. Let us see if this does the trick for us:

$$\begin{aligned} |2\rangle &= \frac{1}{\sqrt{2}}\left(|+-\rangle + |-+\rangle\right) \\ |3\rangle &= \frac{1}{\sqrt{2}}\left(|+-\rangle - |-+\rangle\right) \end{aligned}$$

What are we doing here, physically, to the original states? We are rotating the kets in their 2-dimensional space in order to try to diagonalize the middle block of the S^2matrix. We can break down the effect of the total spin operator on the above states and use that to figure out what the matrix elements will look like. For instance:

$$S^2 = S_1^2 + S_2^2 + 2S_{1z}S_{2z} + S_{1+}S_{2-} + S_{1-}S_{2+}.$$

We can then consider the activity of each piece on these kets:

$$\begin{aligned} S_1^2\,|2\rangle &= \left[\frac{1}{2}(\frac{1}{2}+1)\hbar^2 + \frac{1}{2}(\frac{1}{2}+1)\hbar^2\right]|2\rangle = \frac{3}{2}\hbar^2\,|2\rangle \\ S_1^2\,|3\rangle &= \left[\frac{1}{2}(\frac{1}{2}+1)\hbar^2 - \frac{1}{2}(\frac{1}{2}+1)\hbar^2\right]|3\rangle = 0\hbar^2\,|3\rangle \end{aligned}$$

Similarly,

$$\begin{aligned} S_2^2\,|2\rangle &= \frac{3}{2}\hbar^2\,|2\rangle \\ S_2^2\,|3\rangle &= 0\hbar^2\,|3\rangle \end{aligned}$$

Consider then:

$$\begin{aligned} S_{1z}S_{2z}\,|2\rangle &= -\frac{1}{4}\hbar^2\,|2\rangle \\ S_{1z}S_{2z}\,|3\rangle &= 0\hbar^2\,|3\rangle \end{aligned}$$

We can then consider the action of the ladder operator cross-terms between the two particles:

$$\begin{aligned} S_{1+}S_{2-}\,|2\rangle &= S_{1+}\left(|0\rangle + C_-(\frac{1}{2},\frac{1}{2})\,|--\rangle\right) \\ &= C_+(\frac{1}{2},-\frac{1}{2})C_-(\frac{1}{2},\frac{1}{2})\,|+-\rangle \end{aligned}$$

and

$$\begin{aligned} S_{1-}S_{2+}\,|2\rangle &= S_{1-}\left(C_+(\frac{1}{2},-\frac{1}{2})\,|++\rangle + |0\rangle\right) \\ &= C_-(\frac{1}{2},\frac{1}{2})C_+(\frac{1}{2},-\frac{1}{2})\,|-+\rangle \end{aligned}$$

so that, together:

$$[S_{1+}S_{2-} + S_{1-}S_{2+}]\,|2\rangle = C_+(\frac{1}{2}, -\frac{1}{2})C_-(\frac{1}{2}, \frac{1}{2})\,|2\rangle = \hbar^2\,|2\rangle$$

while for $|3\rangle$ we can show that:

$$[S_{1+}S_{2-} + S_{1-}S_{2+}]\,|3\rangle = 0\hbar^2\,|3\rangle\,.$$

Putting all the pieces together and applying the S^2 matrix, we find:

$$\begin{aligned}\langle 2|\,S^2\,|2\rangle &= \langle 2|\,\left(2\hbar^2\right)|3\rangle = 2\hbar^2\\ \langle 3|\,S^2\,|3\rangle &= \langle 3|\,0\hbar^2\,|3\rangle = 0\hbar^2\end{aligned}$$

while:

$$\langle 2|\,S^2\,|3\rangle = \langle 3|\,S^2\,|2\rangle = 0\hbar^2$$

What about the S_z operator? Have we maintained its diagonalization? A quick check shows we have - that

$$\begin{aligned}(S_{z1} + S_{z2})\,|2\rangle &= 0\hbar\\ (S_{z1} + S_{z2})\,|3\rangle &= 0\hbar\end{aligned}$$

We conclude that by composing these linear combinations of the single-particle states we have arrived at the eigenvectors of the S^2 and the S_z operator; in this basis, both are diagonal and guaranteed to commute. Three of the eigenvectors have spin-1 (with three different z-projections of the two-particle spins) and one has spin-0. We can denote the states by $\psi_{s\,s_z}$, where s is the total spin quantum number and s_z is the total z-component projection quantum number. We then find:

$$\begin{array}{lll} s=1 & |\psi_{1+}\rangle = & |{+}{+}\rangle \\ & |\psi_{10}\rangle = & \frac{1}{\sqrt{2}}\left(|{+}{-}\rangle + |{-}{+}\rangle\right) \\ & |\psi_{1-}\rangle = & |{-}{-}\rangle \\ \\ s=0 & |\psi_{00}\rangle = & \frac{1}{\sqrt{2}}\left(|{+}{-}\rangle - |{-}{+}\rangle\right) \end{array}$$

6.5 An Application of Addition of Spin Angular Momentum: Meson Spectroscopy

In this brief section, I will show you how useful it is to be able to recognize basic facts about the structure of matter utilizing only the most basic information we derived in the previous section.

Mesons are colorless bound states of a pair of quarks. For a variety of reasons, not the least of which is the kinetic energy available to the quarks in the bound state, the heavier the quark the more we can treat the bound state in a non-relativistic way in QCD (the mathematical theory of the strong interaction). One can imagine a pair of heavy quarks, for instance bottom quarks, bound together in a colorless state. To do this requires one quark and one anti-quark. A simple example is bottomonium, $b\bar{b}$.

Let us consider an orbiting pair, $b\bar{b}$, in a state of $L = 0$ total orbital angular momentum. The only angular momentum available to the system, then, is spin. From the above example, we can hypothesize the existence of a spin-1 "triplet" of states and a spin-0 "singlet" state. The spin-1 triplet consists of three states with differing z-projections of their total spin angular momentum, while the singlet consists only of a single total $S = 0$ state.

Given all the various possible radial orbit configurations, angular momentum configurations, and the spin singlet and tripler structures for each possibility, you get a rich spectrum of states. For the lowest-energy states and for the case where there is no orbital angular momentum (L=0), we simply have the triplet states, 1^3S_1, and the singlet, 1^1S_0. These are known in the HEP community as the "Upsilons" (discovered in 1977), or $\Upsilon(1S)$, and the $\eta_b(1S)$ (the "ay-tuh sub bee" or "ay-tuh bee"), respectively. The $\eta_b(1S)$ was only discovered in 2008, first by the BaBar experiment and then confirmed by a second BaBar Collaboration measurement and then an independent measurement by the Belle Collaboration.

7 The General Problem of Adding Angular Momentum

In general, our system of particles may contain those with varying spin angular momenta and also orbital angular momentum. We need to appeal to the notation developed in Section 6.2 and proceed to outline the general problem of adding angular momentum of both sorts.

When the J_z operator acts on a state described by the ket of a two-particle system (involving both orbital and spin angular momentum, we expect to find that:

$$J_z \left| j_1 m_1, j_2 m_2 \right\rangle = (m_1 + m_2)\hbar \left| j_1 m_1, j_2 m_2 \right\rangle .$$

We know already from some experience with the two-spin-$\frac{1}{2}$ particle case that the matrix of this operator will be both diagonal and degenerate (there are multiple ways to combine the z-components of the two particles' spins and orbital angular momenta and still achieve the same total z-projection). The exception to the previous statement is when $m = \pm(j_1 + j_2)$; in those cases, there is just one way to achieve each of the two possible maximum projections.

The general problem of adding angular momentum involves:

- Recognizing that the kets representing the allowed ways of orienting L and S are not necessarily the eigenstates of both $J^2$and J_z.
- Solving the eigenvalue problem to identify the common eigenstates of the two operators (which, of course, means making J_x and J_y non-diagonal).
- By doing this, we obtain the common eigenstates for the two operators and thus the measurable states of the system. This is really what is meant by "solving the general problem of adding angular momentum" - *identifying the physical, measureable states of the multi-particle system.*

We can ask a few questions now.

1. How many kets will there be? For an n-particle system, the total number of states is given by:

$$\prod_{i=1}^{n} (2j_i + 1)$$

   For example, in a system of 2 particles with $\ell = 0$ and $s = \frac{1}{2}$ (the case we did in the notes earlier), we expect $(2s_1+1)(2s_2+1) = 4$ total kets. This is, in fact, what we found. For the problem in Homework 3 involving $\ell_i = 0$ and $s_i = \frac{1}{2}$ for three particles, we expect $(2s_1+1)(2s_2+1)(2s_3+1) = 8$ total states (which, again, you can verify manually . . . but this makes it so much simpler to find the total number of kets).

2. How will we label the states? The states can be labeled by their total angular momentum, the z-projection of the total angular momentum, and the total angular momenta of each individual particle in the space:

$$\left| jm, j_1 j_2 \right\rangle \text{ with } j_1 + j_2 \geq j \geq j_1 - j_2,\ j \geq m \geq -j.$$

   This can be seen by thinking about what happens when you add two simple space vectors. The maximum-length vector you can make from the two is one whose length is the sum of each of their individual lengths; the shortest you can make has a length given by the difference of their individual lengths. The z-projections for each j state simply go from j to $-j$. Shankar writes the possible kets in matrix format, labeling each row and column by kets of total j, $|j, m\rangle$, suppressing the $j_1$and j_2 labels

to simplify the notation:

$$\begin{array}{ccccc}
|j_1+j_2, j_1+j_2\rangle & & & & \\
|j_1+j_2, j_1+j_2-1\rangle & |j_1+j_2-1, j_1+j_2-1\rangle & & & \\
|j_1+j_2, j_1+j_2-2\rangle & |j_1+j_2-1, j_1+j_2-2\rangle & & |j_1-j_2, j_1-j_2\rangle & \\
. & . & & . & \\
. & . & . \quad . \quad . & . & \\
. & . & & . & \\
|j_1+j_2, -(, j_1+j_2-2)\rangle & |j_1+j_2-1, -(, j_1+j_2-2)\rangle & & |j_1-j_2, -(j_1-j_2)\rangle & \\
|j_1+j_2, -(, j_1+j_2-1)\rangle & |j_1+j_2-1, -(, j_1+j_2-1)\rangle & & & \\
|j_1+j_2, -(, j_1+j_2)\rangle & & & &
\end{array}$$

3. How does one then express the kets of total-j in terms of *linear combinations of the product kets?* That is, how do you relate the above states to the states $|j_1 m_1, j_2 m_2\rangle$? That is the hardest part, as you learned on Homework 3.

7.1 Application to the simple system of two spin-$1/2$ particles

To answer the last question, let us revisit the two-spin-$\frac{1}{2}$ particle example from earlier in the notes. Rather than grinding through all that matrix algebra, we could instead have started from the total-j kets we expect for this system: $j_{max} = j_1 + j_2 = s_1 + s_2 = 1$, and $j_{min} = |j_1 - j_2| = |s_1 - s_2| = 0$. There are only two possibilities. So:

$$\begin{array}{cc}
|1,1\rangle & \\
|1,0\rangle & |0,0\rangle \\
|1,-1\rangle &
\end{array}$$

Consider the TOP STATE in each column. For the first column, this is simply:

$$|1,1\rangle = |{+}{+}\rangle\,.$$

As we said earlier, there is only one way to put both spin projections up, so there is an easy identity between the total-j ket and the product ket, given above. How do we then get the other states in this column?

Simple: apply the total angular momentum lowering operator - let it do all the work for you. In this case:

$$J_- |1,1\rangle = J_- |{+}{+}\rangle$$

is what we want to do next. We know that:

$$J_- |1,1\rangle = C_-(1,1)\,|1,0\rangle$$

by definition. The coefficient we know how to calculate:

$$J_- |j,m\rangle = \hbar\sqrt{(j+m)(j-m+1)}\,|j,m-1\rangle\,.$$

In this case:

$$j = 1, m = 1:\ C_-(1,1) = \hbar\sqrt{2}.$$

So we have:

$$J_- |1,1\rangle = \hbar\sqrt{2}\,|1,0\rangle\,.$$

We then need to figure out how this then relates to the product kets. We can also write:

$$\begin{aligned} J_- |++\rangle &= (J_{1-} + J_{2-}) |++\rangle \\ &= C_-(\frac{1}{2}, \frac{1}{2}) |-+\rangle + C_-(\frac{1}{2}, \frac{1}{2}) |+-\rangle \\ &= \hbar |-+\rangle + \hbar |+-\rangle . \end{aligned}$$

Combining what we have learned:

$$J_- |1,1\rangle = \hbar\sqrt{2} |1,0\rangle = \hbar (|+-\rangle + |-+\rangle)$$

and solving for the state we want:

$$|1,0\rangle = \frac{1}{\sqrt{2}} (|+-\rangle + |-+\rangle)$$

which is EXACTLY what we found from all sorts of complicated matrix computation earlier! This is much simpler! If one lowers again, one finds what one expects:

$$|1,-1\rangle = |--\rangle .$$

Now, what about the $j = 0$ state? Again, it will be a linear combination of the only two $m = 0$ states that are available: $|+-\rangle$ and $|-+\rangle$. But, it won't be the same linear combination that yielded $|1,0\rangle$. It must be ORTHOGONAL to that linear combination, and its coefficients must be real (as a matter of convention - the complex portions are absorbed into the kets themselves, by convention). Thus:

$$|0,0\rangle = \alpha |+-\rangle + \beta |-+\rangle .$$

Applying these constraints, we find:

1. From orthogonality:

$$\begin{aligned} \langle 0,0|1,0\rangle &= 0 = \frac{1}{\sqrt{2}} (\alpha + \beta) \\ 0 &= \alpha + \beta. \end{aligned}$$

2. From the constraint of real coefficients:

$$\langle 0,0|0,0\rangle = \alpha^2 + \beta^2 = 1.$$

If we then solve, we find:

$$\alpha = -\beta$$

and so we can choose $\alpha = 1$ and $\beta = -1$, yielding:

$$|0,0\rangle = \frac{1}{\sqrt{2}} (|+-\rangle - |-+\rangle) .$$

Again, no need to go through all the messy diagonalization. Things get a lot simpler here.

This is the problem of finding the Clebsch-Gordon Coefficients - the numbers that multiple the states in the linear combination of product kets needed to express the total-j kets (more on this problem generally in a moment). To find the top state in the next column (after to finish the easy one), and based on the constraints of real coefficients and orthonormality, we find:

$$|j_1 + j_2 - 1, j_1 + j_2 - 1\rangle = \left(\frac{j_1}{j_1 + j_2}\right)^{1/2} |j_1 j_1, j_2 (j_2 - 1)\rangle - \left(\frac{j_2}{j_1 + j_2}\right)^{1/2} |j_1 (j_1 - 1), j_2 j_2\rangle$$

for the two-particle case.

7.2 The Clebsch-Gordon Coefficients

The problem we are trying to solve, in general, boils down to finding those messy coefficients that one needs in order to combine the product kets with the total-j kets. That's it. That's the problem.

These coefficients are called *Clebsch-Gordon Coefficients*. We can write them in general bra-ket notation as:

$$|jm, j_1 j_2\rangle = \sum_{m_1} \sum_{m_2} |j_1 m_1, j_2 m_2\rangle \langle j_1 m_1, j_2 m_2 | jm, j_1 j_2\rangle$$

where

$$\langle j_1 m_1, j_2 m_2, ..., j_n m_n | jm, j_1 j_2 ... j_n\rangle = \langle j_1 m_1, j_2 m_2, ..., j_n m_n | jm\rangle$$

are the Clebsch-Gordon Coefficients themselves. They have the following properties:

1. $\langle j_1 m_1, j_2 m_2 | jm\rangle \neq 0$ only if $|j_1 - j_2| \leq j \leq j_1 + j_2$ - this is the implication of the Triangle Inequality. We must be able to form a triangle with sides of length j_1, j_2, and j .
2. $\langle j_1 m_1, j_2 m_2 | jm\rangle \neq 0$ only if $m_1 + m_2 = m$
3. They are real, by convention
4. $\langle j_1 j_1, j_2 (j - j_1 | jj\rangle > 0$ by convention (this fixes the sign of the top state)
5. $\langle j_1 m_1, j_2 m_2 | jm\rangle = (-1)^{j_1 + j_2 - j} \langle j_1 (-m_1), j_2 (-m_2) | j(-m)\rangle$ - this tells us the coefficients for our negative m states given our positive m states.

7.3 The Explicit Formula for Clebsch-Gordon Coefficients

The explicit formula for Clebsch-Gordon Coefficiencts is given by:

$$\begin{aligned} C^{m_1, m_2, m}_{j_1, j_2, j} &= \delta_{m, m_1 + m_2} \sqrt{\frac{(2j+1)(j + j_1 - j_2)!(j - j_1 + j_2)!(j_1 + j_2 + j)}{(j_1 + j_2 + j + 1)!}} \\ &\times \sqrt{(j+m)!(j-m)!(j_1 - m_1)!(j_1 + m_1)!(j_2 - m_2)!(j_2 + m_2)!} \\ &\times \sum_k \frac{(-1)^k}{k!(j_1 + j_1 - j - k)!(j_1 - m_1 - k)!(j_2 + m_2 - k)!(j - j_2 + m_1 - k)!(j - j_1 - m_2 + k)!} \end{aligned}$$

where k is any zero or positive integer such that the factorial argument is non-negative.

The Particle Data Guide [25] contains a helpful table of these coefficients, reproduced in Fig. 6.

42. CLEBSCH-GORDAN COEFFICIENTS, SPHERICAL HARMONICS, AND d FUNCTIONS

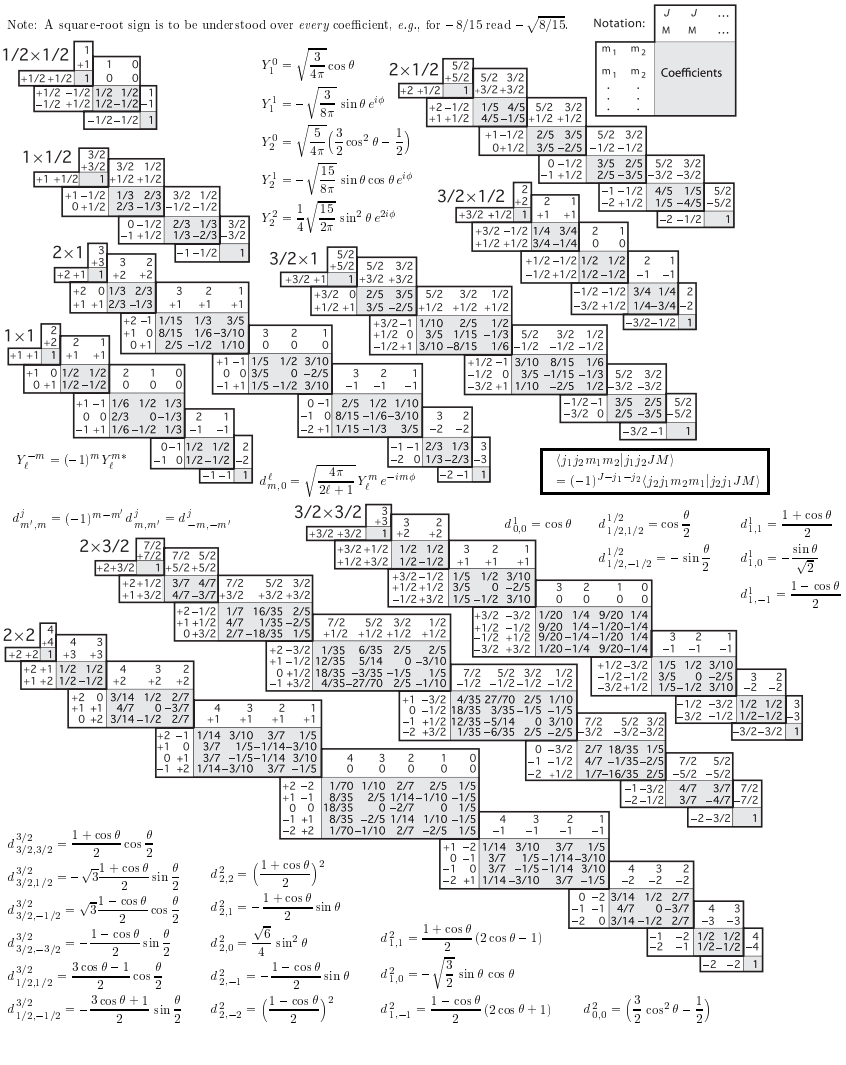


Figure 42.1: The sign convention is that of Wigner (*Group Theory*, Academic Press, New York, 1959), also used by Condon and Shortley (*The Theory of Atomic Spectra*, Cambridge Univ. Press, New York, 1953), Rose (*Elementary Theory of Angular Momentum*, Wiley, New York, 1957), and Cohen (*Tables of the Clebsch-Gordan Coefficients*, North American Rockwell Science Center, Thousand Oaks, Calif., 1974).

Figure 6: Reproduction of the table of Clebsch-Gordon Coefficients from Ref. [25].

7.4 Comment: Application to the three spin-1/2 particle system

Let's try applying this to the more complicated problem in Homework 3: three spin-$\frac{1}{2}$ particles. We can only begin the computation, as the second major step involves a more general version of the Clebsch-Gordon Coefficients.

The highest state we can get is $|+++\rangle$. This state has $j_{max} = j_1 + j_2 + j_3 = \frac{3}{2}$ and the maximum z-projection. Thus we identify:

$$\left|\frac{3}{2}, \frac{3}{2}\right\rangle = |+++\rangle\,.$$

We then apply the lowering operator to find the next state in this column:

$$\begin{aligned} J_- \,|3/2, 3/2\rangle &= \hbar\sqrt{(3/2+3/2)(3/2-3/2+1)}\,|3/2, 1/2\rangle \\ &= \hbar\sqrt{3}\,|3/2, 1/2\rangle\,. \end{aligned}$$

Also,

$$\begin{aligned} J_- \,|+++\rangle &= (J_{1-} + J_{2-} + J_{3-})\,|+++\rangle \\ &= \hbar\,|-++\rangle + \hbar\,|+-+\rangle + \hbar\,|++-\rangle\,. \end{aligned}$$

Combining this, we learn

$$|3/2, 1/2\rangle = \frac{1}{\sqrt{3}}\left(|-++\rangle + |+-+\rangle + |++-\rangle\right).$$

Continuing, we confirm the other states:

$$\begin{aligned} |3/2, -1/2\rangle &= \frac{1}{\sqrt{3}}\left(|--+\rangle + |-+-\rangle + |+--\rangle\right) \\ |3/2, -3/2\rangle &= |---\rangle \end{aligned}$$

How do we then get to the next columns of states? Well, let's start with finding j_{min}. This is determined from:

$$j_{min} = |j_1 + j_2 - j_3| = |j_1 - j_2 + j_3|$$

which tells us that there are two ways to form the minimum state by combining individual j_i quantum numbers. Well, in that column we know that we have the top-most state with $m = j_1 + j_2 + j_3 - 1 = 1/2$. We know that it will be orthogonal to $|3/2, 1/2\rangle$ but will have the same z-projection. It must be normalized to unity. The product kets that yield $m = 1/2$ are:

$$|++-\rangle\,,\ |+-+\rangle\ |-++\rangle$$

We then have to determine the highest state in the next column. The Clebsch-Gordon Coefficients are determined for the case where 2 particles combine into single-particle hybrid states. We have a 3-particle case. We need a more general version of the Clebsch-Gordon Coefficients.

I won't go through the rest of the calculation, but merely point the way. We need the *Wigner 3-jm symbols* to go to a 3-particle system. They are written as:

$$|jm, j_1 j_2 j_3\rangle = \sum_{m_1}\sum_{m_2}\sum_{m_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} |j_1 m_1, j_2 m_2, j_3 m_3\rangle\,.$$

Many computational frameworks, like Mathematica, provide functions to calculate these for you. They are related to the Spherical Harmonics in that they give you the integral of the product of three spherical harmonic functions.

7.5 Particle Decay

The addition of angular momentum is useful not only for finding the structure of bound states of particles; you can also use it to determine the possible decay modes of a particle, based purely on restrictions on angular momentum imposed by conservation. For instance, we can explore how the spin-0 Higgs Boson might decay (while at rest) to various final states. We want to find out how to express the state $|0,0\rangle$ in terms of product kets for various possible final states, such as 2 spin-1/2 particles, two spin-1 particles, etc.

We can begin with the two spin-1/2 particle case, and use a table of Clebsch-Gordon Coefficients to figure out the answer. We know we want the total-j ket, $|0,0\rangle$, and we want it to decay to a pair of spin-1/2 particles. For instance, how are the angular momenta of a pair of tau leptons, or bottom quarks, arranged in product ket space when they are produced from a parent spin-0 particle? Reading from the table for such a pair,

$$|0,0\rangle = \frac{1}{\sqrt{2}}\left(\left|\frac{1}{2},+\frac{1}{2}\right\rangle\left|\frac{1}{2},-\frac{1}{2}\right\rangle - \left|\frac{1}{2},-\frac{1}{2}\right\rangle\left|\frac{1}{2},+\frac{1}{2}\right\rangle\right).$$

We've already seen this, of course, when trying to compose a new state out of an original pair of spin-1/2 particles.

How about the decay of the Higgs boson to something like a pair of Z or W bosons, which each have spin-1? What might this look like? Reading from the table for 1,1 Clebsch-Gordon Coefficients:

$$|0,0\rangle = \frac{1}{\sqrt{3}}\left(|1,+1\rangle|1,-1\rangle - |1,0\rangle|1,0\rangle + |1,-1\rangle|1,+1\rangle\right)$$

We see that such a decay is possible, and contains a complicated (but apparently uniformly probable) distribution of polarizations of the spins of the spin-1 particles in the final state.

Consider the decay of a Z-boson to a pair of fermions with spin-1/2. In that case, we have a $S = 1$ particle decaying into two $S = 1/2$ particles. We can write down the relationship between the total-j kets that describe the possible states of the Z boson and the product kets that describe the possible states of the individual particles in the final state,

$$\begin{aligned}
|1,1\rangle &= \left|\frac{1}{2},\frac{1}{2}\right\rangle\left|\frac{1}{2},\frac{1}{2}\right\rangle \\
|1,0\rangle &= \frac{1}{\sqrt{2}}\left(\left|\frac{1}{2},\frac{1}{2}\right\rangle\left|\frac{1}{2},-\frac{1}{2}\right\rangle + \left|\frac{1}{2},-\frac{1}{2}\right\rangle\left|\frac{1}{2},\frac{1}{2}\right\rangle\right) \\
|1,-1\rangle &= \left|\frac{1}{2},-\frac{1}{2}\right\rangle\left|\frac{1}{2},-\frac{1}{2}\right\rangle
\end{aligned}$$

We see that there is a rich set of structure in the outcomes of such a decay. We'll explore this further in the future.

8 Relativistic Quantum Mechanics

For the rest of these notes, we will use a standard set of units common to high-energy particle physics named "natural units." Since we are typically dealing with speeds near that of light and angular momenta that are at the subatomic level, we will adopt the following convention:

$$\begin{aligned}
\hbar &= 1 \\
c &= 1
\end{aligned}$$

All speeds and angular momenta are normalized to that of light and the reduced Planck's Constant, respectively. Since:

$$\hbar c \approx 197\,\text{MeV}\cdot\text{fm},$$

we can relate energy and distance through Planck's constant and the speed of light. Namely:

$$\hbar c = 1 \longrightarrow 1\,\text{fm} = \frac{1}{197}\,\text{MeV}^{-1}.$$

In this system of units, the Schroedinger Wave Equation for a free particle would be written:

$$\begin{aligned} H\,|\psi\rangle &= \frac{1}{2m}P^2\,|\psi\rangle \\ i\hbar\frac{\partial}{\partial t}\,|\psi\rangle &= -\frac{\hbar^2}{2m}\nabla^2\,|\psi\rangle \\ i\frac{\partial}{\partial t}\,|\psi\rangle &= -\frac{1}{2m}\nabla^2\,|\psi\rangle \end{aligned}$$

again, keeping in mind that $E = i\frac{\partial}{\partial t}$ and $\vec{P} = i\hbar\vec{\nabla}$.

Remember also that the conservation of energy equation looks as follows:

$$E^2 = p^2c^2 + m^2c^4,$$

which simplifies in natural units to:

$$E^2 = p^2 + m^2.$$

The invariant of the above equation is the mass, and so expressing this as an equation in terms of the invariant and the variable components (total energy and momentum):

$$m^2 = E^2 - p^2.$$

We can express energy and momentum as a four-vector:

$$\mathbf{p} = (E, \vec{p}) \equiv p^\mu.$$

To compute the square of this four-vector and recover the invariant:

$$\mathbf{p}^2 = m^2,$$

we need to introduce a matrix in between the product of the two vectors that introduces the sign flip. We can see what this must look like:

$$m^2 = \mathbf{p}^2 = \begin{bmatrix} E & p_x & p_y & p_z \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} E \\ p_x \\ p_y \\ p_z \end{bmatrix} = E^2 - p^2.$$

We refer to p_μ (the column vector, in this case) as the contravariant vector and $p^\nu = \sum_\nu p_\nu g^{\mu\nu}$ as the covariant vector, the product of the row vector with this matrix, $g^{\mu\nu}$. This new 4x4 matrix is known as "the metric" - it transforms covariant into contravariant vectors (and vice versa). If you want to learn more about all of this, dig a bit into general relativity (which relies on this concept entirely). Contravariant vectors transform in the same way as the coordinates in a coordinate system, while covariant vectors transform the opposite way of the coordinates (the same way that the coordinate axes would change - they "co-vary" with the axes under a transformation of the coordinate system).

8.1 The Dirac Equation

The Schroedinger Wave Equation is an incomplete description of nature in several ways:

- It does not incorporate the postulates of special relativity (for instance, its expression of energy conservation is explicitly classical, ignoring internal energy).
- It failes to recognize that the potentials themselves may be quantized (as one might expect since photons are invisible units that transmit the electromagnetic force, yet nowhere does the photon appear in the SWE).

We can resolve the first but not the second in this lecture series. Resolving the first will already be sufficient to come to a complete, quantum understanding of spin. Spin has so far been an ad hoc two-state phenomenon that we have added by hand in the SWE. We will now step back to first principles (energy conservation and the special theory of relativity) and derive the correct, relativistic wave equation.

Paul Adrian Maurice Dirac derived this equation and his complete work was published in 1932 [20]. However, the derivation of the Dirac Equation occurred in 1928. We will derive this equation and determine the form of its solutions, considering a "free particle" case (as Dirac did).

The special relativistic expression of energy conservation is not a linear equation; it is a quadratic equation:

$$E^2 = m^2 + |\vec{p}|^2$$

which admits both positive and negative energy solutions. Dirac sought a LINEAR expression of the same form, to avoid this problem. That linear expression would then serve the role of the basis of a relativistics wave equation, akin to the SWE:

$$H\,|\psi\rangle = E\,|\psi\rangle \text{ (free particle case).}$$

We can begin with the hypothesis that:

$$H\,|\psi\rangle = \left(\vec{\alpha}\cdot\vec{P} + \beta m\right)|\psi\rangle\,,$$

which is a guess at a linear equation that now includes internal energy (mass). We must be able to recover from this the special relativistic expression of energy conservation (which is the correct expression, after all). Thus we must find that:

$$H^2\,|\psi\rangle = \left(\vec{P}^2 + m^2\right)|\psi\rangle\,.$$

These two equations together represent the Dirac Equation; both must hold to be true, in order to marry the SWE with relativity.

We can press forward and determine how this equation describes particles absent external influences. We need to determine these unknown coefficients - $\vec{\alpha}$and β. Algebraically:

$$H^2\,|\psi\rangle = (\vec{\alpha}\cdot\vec{P} + \beta m)(\vec{\alpha}\cdot\vec{P} + \beta m)\,|\psi\rangle$$

It is convenient to employ Einstein Summation Notation for the continued working of this equation. In Einstein Summation Notation, any repeated index that appears in an equation represents and implicit sum over that index. So, instead of writing:

$$\vec{\alpha}\cdot\vec{P} = \sum_{i=0}^{3}\alpha_i P_i$$

we write

$$\vec{\alpha}\cdot\vec{P} = \alpha_i P_i$$

and imply that the sum must be taken over all values of i. If we adopt this notation, we can resume in a more simple and approachable way the squaring of terms we have just undertaken:

$$\begin{aligned} H^2 |\psi\rangle &= (\alpha_i P_i + \beta m)(\alpha_j P_j + \beta m) |\psi\rangle \\ &= \left(\alpha_i^2 P_i^2 + (\alpha_i \alpha_j + \alpha_j \alpha_i) P_i P_j + (\alpha_i \beta + \beta \alpha_i) P_i m + \beta^2 m^2\right) |\psi\rangle . \end{aligned}$$

We already see something quite interesting here, writing out all the unique terms:

- The unknown coefficients cannot be numbers. Why? Because in order for the above equation to match the relativistic expecation,

  $$H^2 |\psi\rangle = (P^2 + m^2) |\psi\rangle ,$$

  it must be true that $\alpha_i \alpha_j + \alpha_j \alpha_i = 0$ and $\alpha_i \beta + \beta \alpha_i = 0$. These coefficiencts must anti-commute. Numbers don't anti-commute. These objects must be, at minimum, *matrices* (or, in a general language, *tensors*).

- The square of these coefficients must yield the identity:

  $$\alpha_i^2 = I \text{ and } \beta^2 = I.$$

  Again, this is to satisfy the requirement that squaring the Hamiltonian must yield and expression for energy conservation consistent with special relativity.

We have certainly encountered objects that anti-commute like this, with these rules: the Pauli Spin Matrices. You explored some of these properties on your first homework. That effort was not in vain.

8.1.1 The Coefficients of the Dirac Equation

We are left to determine the exact form of these coefficients, given the above constraints from special relativity. In fact, going a bit further, one will find that these matrices must satisfy the following additional constraints:

- They must be Hermitian.
- They must be traceless.
- They must be of even dimensionality (2x2, 4x4, 6x6, etc.)
- They must have eigenvalues of ± 1

The minimum dimension matrices that satisfy all four of these requirements are 4x4. The choice of the matric representation is not unique, but we will employ the Dirac-Pauli representation:

$$\vec{\alpha} = \begin{bmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{bmatrix} \quad \beta = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} .$$

The physics should only depend on the properties of the matrices and not their specific representations.

8.1.2 The Solutions to the Dirac Equation

We see already that the solutions to this equation, $|\psi\rangle$, must be represented in matrix notation minimally by a 4-row column-vector. This solution is referred to as a *Dirac Spinor*. We have four solutions for each "particle" for this equation . . . a bounty of solutions, whose physical meaning needs to be understood.

8.1.3 The Covariant Form of the Dirac Equation

We have arrived at a form for the Dirac Equation:

$$H\left|\psi\right\rangle = i\hbar\frac{\partial}{\partial t}\left|\psi\right\rangle = \left(\left[\begin{array}{cc} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{array}\right]\cdot i\hbar + \left[\begin{array}{cc} I & 0 \\ 0 & -I \end{array}\right]m\right)\left|\psi\right\rangle.$$

It is useful to write this in another form. Multiplying from the left by β:

$$\begin{aligned} i\beta\frac{\partial}{\partial t}\left|\psi\right\rangle &= -i\beta\vec{\alpha}\cdot\vec{\nabla}\left|\psi\right\rangle + m\left|\psi\right\rangle \\ \left[i\left(\beta\frac{\partial}{\partial t} + \beta\vec{\alpha}\cdot\vec{\nabla}\right) - m\right]\left|\psi\right\rangle &= 0 \end{aligned}$$

The operator on the left-most side of the above equation looks like the product of two four-vectors. One of the four-vectors is one containing the partial derivatives over space and time:

$$\partial_\mu = \left(\frac{\partial}{\partial t}, \vec{\nabla}\right)$$

while the other is a four-vector of matrices is

$$\gamma^\mu = (\beta, \beta\vec{\alpha})$$

The above form is known as the covariant form of the Dirac Equation, and can be simply written as:

$$\left[i\gamma^\mu\partial_\mu - m\right]\left|\psi\right\rangle = 0.$$

The solutions to this equation will be the wave functions we seek. There are four components to the wave function; this equation can be understood to be a set of FOUR differential equations that couple the four components of a single column vector. This is most easily seen by writing this as:

$$\sum_{k=1}^{4}\left[\sum_\mu i\left(\gamma^\mu\right)_{jk}\partial_\mu - m\delta_{jk}\right]\left|\psi_k\right\rangle = 0.$$

8.2 Solutions of the Dirac Equation

We can now proceed to solve the eigenvector and eigenvalue problem of the Dirac Equation. We can guess at the form of the solutions, which will look much like the old plane-wave solutions of the SWE but now with an unknown 4-component column matrix attached:

$$\left|\psi\right\rangle = u(p)e^{-ip\cdot x}$$

where p is a four-vector, $p = (E, p_x, p_y, p_z)$, as is $x = (t, x, y, z)$. The object $u(p)$ is the unknown four-component spinor whose form we need to determine. If we substitute this solution form into the Dirac Equation:

$$\begin{aligned} \left[i\gamma^\mu\partial_\mu - m\right]u(p)e^{-ip\cdot x} &= 0 \\ \left[i\gamma^\mu(\partial_\mu(u(p)e^{-ip\cdot x})) - m\,u(p)e^{-ip\cdot x}\right] &= 0 \\ \left[i\gamma^\mu u(p)(\partial_\mu(-ip\cdot x))e^{-ip\cdot x} - m\,u(p)e^{-ip\cdot x}\right] &= 0 \\ \left[i\gamma^\mu u(p)(-ip_\mu)e^{-ip\cdot x} - m\,u(p)e^{-ip\cdot x}\right] &= 0 \\ \left[i\gamma^\mu(-ip_\mu)\,u(p)e^{-ip\cdot x} - m\,u(p)e^{-ip\cdot x}\right] &= 0 \\ \left[\gamma^\mu p_\mu - m\right]u(p) &= 0 \end{aligned}$$

It is common to denote covariant products of gamma matrices with other matrices or vectors as:

$$\not{p} = \gamma^\mu p_\mu.$$

We can then write:

$$[\not{p} - m]\, u(p) = 0.$$

We now want to find the solutions to this. This is just an eigenvalue equation, so we already know that the eigenvalues are $\pm m$. They are masses. We need to identify the eigenvectors that go with these eigenvalues. We can see this if we take the momentum of the particle to be zero (at rest). Then $p = 0$ and

$$\begin{aligned} H\, u(p) &= (\vec{\alpha} \cdot \vec{p} + \beta m) u(p) = E u(p) \\ H\, u(p) &= \beta m\, u(p) = \begin{bmatrix} mI & 0 \\ 0 & -mI \end{bmatrix} u(p). \end{aligned}$$

This equation yields four eigenvalues: m, m, $-m$, and $-m$. That means we have two positive-energy states and two negative energy states. The eigenvectors can be seen to be:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

The Feynman-Stuckelburg interpretation of these states is the one commonly adopted in modern practice. In this interpretation, we ascribe the first two eigenstates with E>0 to be the particle eigenstates. The second pair of eigenstates as E>0 antiu-particle states, to avoid the problem of negative energies.

If we allow for non-negative momentum, then the eigenvalue equation merely becomes:

$$H\, u(p) = \begin{bmatrix} mI & \vec{\sigma} \cdot \vec{p} \\ \vec{\sigma} \cdot \vec{p} & -mI \end{bmatrix} \begin{bmatrix} u_A \\ u_B \end{bmatrix} = E \begin{bmatrix} u_A \\ u_B \end{bmatrix}.$$

Here, we have split the 4-component column vector (the overall Dirac Spinor) into two, two-component spinors. We then see that:

$$\begin{aligned} \vec{\sigma} \cdot \vec{p}\, u_B &= (E - m) u_A \\ \vec{\sigma} \cdot \vec{p}\, u_A &= (E + m) u_B \end{aligned}$$

For the two E>0 solutions, we can take:

$$u_A^{(s)} = \chi^{(s)}$$

where

$$\chi^{(1)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ \chi^{(2)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

We then have only to specify the lower components of the four-component eigenvectors by inserting these choices into the second equation above:

$$\vec{\sigma} \cdot \vec{p}\, \chi^{(s)} = (E + m) u_B^{(s)}.$$

Solving for $u_B^{(s)}$ we find:

$$u_B^{(s)} = \frac{\vec{\sigma} \cdot \vec{p}}{E + m} \chi^{(s)}.$$

The positive-energy solutions are then:

$$u^{(s)} = N \begin{pmatrix} \chi^{(s)} \\ \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi^{(s)} \end{pmatrix}, \qquad E > 0.$$

Here, N is the normalization constant.

We can repeat this procedure for the E<0 solutions. Here, we can take:

$$u_B^{(s)} = \chi^{(s)}.$$

Then:

$$u_A^{(s)} = \frac{\vec{\sigma}\cdot\vec{p}}{E-m}u_B^{(s)} = -\frac{\vec{\sigma}\cdot\vec{p}}{|E|+m}\chi^{(s)}.$$

Thus the negative energy solutions are:

$$u^{(s+2)} = N \begin{pmatrix} -\frac{\vec{\sigma}\cdot\vec{p}}{|E|+m}\chi^{(s)} \\ \chi^{(s)} \end{pmatrix}, \qquad E < 0.$$

8.3 Spin and the Dirac Equation

We now see where spin comes from. Enforcing special relativity as the correct description of nature at all velocities, we are required to write a wave equation (the Dirac Equation) that is linear in energy and momentum but solved minimally only by four-component wave functions. The solutions above tell us the rest of the story. They are each four-component objects, and so for a single particle there are FOUR solutions to the equation: two are positive-energy, and two are negative-energy (interpreted to mean E>0 anti-particle states). But we see that for the particle or the anti-particle, there are still TWO solutions. There is an EXTRA two-fold degeneracy for each particle which was not present in the non-relativistic SWE. Spin is an inevitable consequence of a universe that obeys the postulates of special relativity. We are forced to have it. This is amazing.

This also implies something else. Since there is an extra two-fold degeneracy in the solutions, there must also be one more observable in nature that commutes with both the Hamiltonian and the momentum operator. The eigenvalues of this additional operator can be used to distinguish the states. Just as in the case of adding angular momentum and having degeneracy in the S_z matrix which is resolved by diagonalizing the S^2 matrix, we have a situation where there is a two-fold degeneracy left in the problem even after working through the H and P parts of the problem.

A common choice for this additional observable is the following, which you can show commutes with H and P:

$$\vec{\Sigma}\cdot\hat{p} = \begin{pmatrix} \vec{\sigma}\cdot\hat{p} & 0 \\ 0 & \vec{\sigma}\cdot\hat{p} \end{pmatrix},$$

where $\hat{p} = \vec{p}/|\vec{p}|$ is a unit vector pointing in the direction of momentum. We can multiply this by any constant we like and preserve the properties of the original matrix. Therefore, we can choose to consider:

$$\vec{S} = \frac{1}{2}\vec{\sigma}$$

and we can think of this new observable as the projection of spin along the direction of motion of the particle:

$$\frac{1}{2}\vec{\sigma}\cdot\hat{p}.$$

This projection is known as "helicity." We can see that this projection has TWO possible eigenvalues:

$$\lambda = \begin{cases} +\frac{1}{2} & \text{"positive helicity"} \\ -\frac{1}{2} & \text{"negative helicity"} \end{cases}$$

9 Spin and Particle Decay

In this section, we will explore how angular momentum (including spin) affects the angular distributions of final-state particles after an initial state decays. I am indebted to Kent Hornbostel for his insights that led to the writing of this part of the notes. This is a synthesis of what we have developed so far.

We will begin by understanding the relationship between angular momentum and rotation. We will then consider what happens if we rotate the final states away from, for instance, a hypothetical z-axis. By considering these two pieces, we will build a toolkit for evaluating angular distributions of final-state particles. This can be further developed when considering the full helicity formalism for describing particle decay and distributions of final-state particles.

9.1 Angular Momentum and Rotation

We can begin by considering what is means to make a rotation in physical space about a particular axis. Let us choose the axis to be the z-axis. Choosing a point in the x-y plane, marked by (x, y), we can imagine rotating about the z-axis by an angle ϕ. The result is a new coordinate, (x', y'), which can be written in terms of the old coordinate as follows:

$$\begin{aligned} x' &= x\cos\phi - y\sin\phi \\ y' &= x\sin\phi + y\cos\phi. \end{aligned}$$

We see that this is merely a matrix operation on a column vector:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

The 2x2 matrix is a rotation matrix, describing an arbitrary rotation by ϕ about the z-axis. We might write this rotation as:

$$|x'\rangle = U(\phi)\,|x\rangle\,.$$

What if we were interested in a more general problem - how to represent a very small (infinitesimal) rotation of a state around an axis (again, let us choose the z-axis)? For a zero rotation, the rotation matrix is simple:

$$U(0) = I,$$

the identity matrix. An infinitestimal rotation, then, would be a small perturbation on top of the identity matrix. We can write the form of this matrix quite simply in terms of the very small rotation angle, denoted as $\delta\phi$,

$$U(\delta\theta) = I - \alpha\,\delta\phi.$$

To get at the form of the matrix added to the identity, we can consider the Taylor Expansion of the cosine and sine functions:

$$\begin{aligned} \sin\delta\phi &= \delta\phi - \frac{1}{3!}\delta\phi^3 + \frac{1}{5!}\delta\phi^5 + ... \\ \cos\delta\phi &= 1 - \frac{1}{2!}\delta\phi^2 + \frac{1}{4!}\delta\phi^4 + ... \end{aligned}$$

and for very small angles (again, infinitesimal rotations)

$$\begin{aligned} \sin\delta\phi &\approx \delta\phi \\ \cos\delta\phi &\approx 1. \end{aligned}$$

Revisiting our rotation of the spatial coordinates:

$$\begin{aligned}
\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} &= \begin{bmatrix} \cos\delta\phi & -\sin\delta\phi & 0 \\ \sin\delta\phi & \cos\delta\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \\
&\approx \begin{bmatrix} 1 & -\delta\phi & 0 \\ \delta\phi & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \\
&= \begin{bmatrix} x \\ y \\ z \end{bmatrix} - \delta\phi \begin{bmatrix} y \\ -x \\ z \end{bmatrix} \\
&= I \begin{bmatrix} x \\ y \\ z \end{bmatrix} - \delta\phi \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}
\end{aligned}$$

This tells us how a pure coordinate state would be rotated, and we can even find the form of α for a pure-coordinate rotation. We might then consider what happens to a physical state (e.g. a function), that depends on these coordinates, when acted upon by the very same matrix. In this case, consider a state

$$|\psi(x,y,z)\rangle\,.$$

We want to determine now the form of the operator that rotates the *wave function* by rotating coordinates about the z-axis. We know that:

$$(I - \delta\phi\,\alpha)\,|x,y,z\rangle = |x - \delta\phi\,y,\ y + \delta\phi\,x\rangle\,.$$

We want to find the transformed wave function, $|\psi'\rangle$, after acting with the rotation operator. We can write:

$$\begin{aligned}
|\psi'\rangle &= (I - \delta\phi\,\alpha)\,|\psi\rangle \\
&= (I - \delta\phi\,\alpha)\int_{-\infty}^{+\infty} |x,y,z\rangle\,\langle x,y,z|\psi\rangle d^3x \\
&= \int_{-\infty}^{+\infty} (I - \delta\phi\,\alpha)\,|x,y,z\rangle\,\langle x,y,z|\psi\rangle d^3x \\
&= \int_{-\infty}^{+\infty} |x - \delta\phi\,y,\ y + \delta\phi\,x,\ z\rangle\,\langle x,y,z|\psi\rangle d^3x \\
&= \int_{-\infty}^{+\infty} |x',y',z'\rangle\,\langle x' + \delta\phi\,y',\ y' - \delta\phi\,x',\ z'|\psi\rangle d^3x'.
\end{aligned}$$

In the last line, we have made the simple variable substitution: $x' = x - \delta\phi y$, $y' = y + \delta\phi x$, and $z' = z$. Multiplying from the left with $\langle x', y', z'|$, only one term in the integral survives the inner product and we find:

$$\langle x',y',z'|\psi'\rangle = \langle x',y',z'|(I - \delta\phi\,\alpha)|\psi\rangle = \langle x' + \delta\phi\,y',\ y' - \delta\phi\,x',\ z'|\psi\rangle.$$

We can identify $\psi(x,y,z) = \langle x,y,z|\psi\rangle$ and $\psi(x + \delta\phi\,y,\ y - \delta\phi\,x,\ z) = \langle x + \delta\phi\,y,\ y - \delta\phi\,x,\ z|\psi\rangle$ (since primes appear on the coordinates on both sides of the above equation, we can merely drop them for convenience).

The last step we need in order to identify our mystery matrix is to simply Taylor Expand the wave function about the small added pieces on each of the x and y coordinates, $y\delta\phi$ and $x\delta\phi$:

$$\langle x,y,z|(I - \delta\phi\,\alpha)|\psi\rangle \quad = \quad \psi(x + \delta\phi\,y,\ y - \delta\phi\,x,\ z)$$

$$
\begin{aligned}
&= \psi(x,y,z) + \frac{\partial\psi}{\partial x}(\delta\phi\, y) + \frac{\partial\psi}{\partial y}(-\delta\phi\, x) + ... \\
&\approx \psi(x,y,z) + \frac{\partial\psi}{\partial x}(\delta\phi\, y) + \frac{\partial\psi}{\partial y}(-\delta\phi\, x) \\
&= \psi(x,y,z) - \delta\phi\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\psi.
\end{aligned}
$$

We are nearly there. Writing out the left side of the above equation and relating the two sides:

$$
\begin{aligned}
\langle x,y,z|(I-\delta\phi\,\alpha)|\psi\rangle &\approx \psi(x,y,z) - \delta\phi\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\psi \\
\psi(x,y,z) - \delta\phi\langle x,y,z|\alpha|\psi\rangle &\approx \psi(x,y,z) - \delta\phi\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\psi.
\end{aligned}
$$

We thus find that:

$$\alpha = \left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$$

in the coordinate basis. This looks suspiciously like

$$L_z = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$$

and so we merely make the identity that:

$$\alpha = \frac{i}{\hbar}L_z.$$

We have arrived at the conclusion that the generator of infinitesimal rotations of the wave function, about the z-axis, is just the L_z angular momentum operator. An infinitesimal rotation about the z-axis is then written:

$$U_z(\delta\phi) = I - \delta\phi\frac{i}{\hbar}L_z.$$

What if we then make a succession of N infinitesimal rotations, each of the same size, $\delta\phi$, such that:

$$N\cdot\delta\phi = \phi,$$

where ϕ is a finite rotation about the z-axis? This would be the same as applying the rotation operator to the state N successive times:

$$\left(I - \frac{\phi}{N}\frac{i}{\hbar}L_z\right)\left(I - \frac{\phi}{N}\frac{i}{\hbar}L_z\right)\left(I - \frac{\phi}{N}\frac{i}{\hbar}L_z\right)...\left(I - \frac{\phi}{N}\frac{i}{\hbar}L_z\right)\psi(x,y,z).$$

In the limit that $N\to\infty$,

$$\lim_{N\to\infty}\left(I - \frac{\phi}{N}\frac{i}{\hbar}L_z\right)^N = e^{-i\phi L_z/\hbar}.$$

We can then write that

$$U_z(\phi) = e^{-i\phi L_z/\hbar}.$$

In general, the rotation matrix in total angular momentum space that rotates one state of $|j,m\rangle$ into another is given by

$$U = e^{-i\theta(\hat{\theta}\cdot\vec{J})/\hbar},$$

where the angle θ is a general angle (this could be decomposed into Euler Angles, for instance, first rotating away from an axis of choice and then rotating around the new direction).

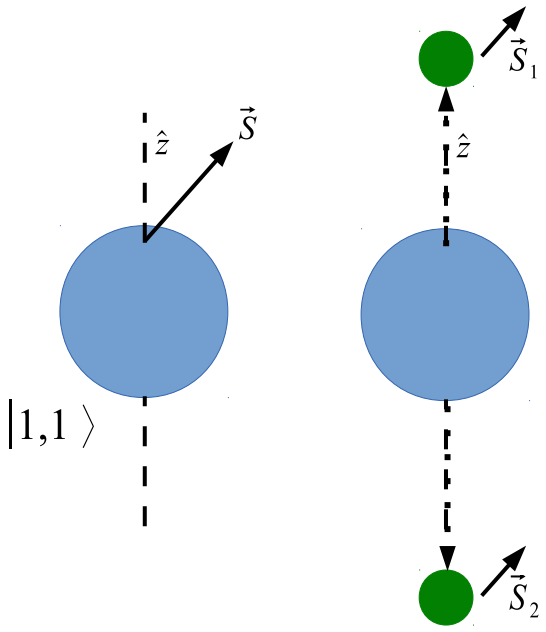


Figure 7: A depiction of the decay of a spin-1 particle (left) into a pair of spin-1/2 particles (right).The z-axis has been conveniently chosen to be the decay axis. The product kets in the final state have positive and negative helicities for particle 1 and particle 2, respectively. Remember - once we have chosen the z-axis as our axis of quantizaton, we cannot know where $\vec{S}$ points; we can only know that the initial state is in a definite z-projection of its spin. The above picture is quite classical, but can aid in thinking.

9.2 Angular Distributions of Final-State Particles

We are now ready to think about how to assemble our pieces:

- We have seen how to write states of total angular momentum in terms of product kets over the individual particles that compose the total state
- We have seen the form that general rotations take, so we are in principle equipped to take a state of total angular momentum ("total-j") and rotate it (and, consequently, the particles that compose it) about some axis.
- We recognize that in fully relativistic quantum mechanics, we have to consider three angular-momentum-related quantum numbers related to internal angular momentum ("spin"): the total angular momentum quantum number, j, the projection along the z-axis, m, and the projection of spin along the direction of motion ("helicity"), λ.

Using these pieces, we can probe the angular distributions of particles resulting from the decay of a total-j state.

9.2.1 Example: the decay of a spin-1 particle into two spin-$\frac{1}{2}$ particles

Consider a spin-1 particle, at rest, that spontaneously decays to a pair of spin-1/2 particles (Fig. 7). Let us prepare it (produce it) such that its spin projects along the positive z-direction with $m = 1$. Let us recall how we wrote this total-j state of spin-1 in terms of the product kets:

$$|1,1\rangle = |{+}{+}\rangle\,.$$

We need to take the helicity quantum number into account, to be fully relativistic (as we learned from the Dirac Equation. We can then label states:

$$|j, m, \lambda\rangle$$

and write the above state of total-j as

$$|1, 1, 1\rangle$$

It is convention to choose, for a particle at rest, the helicity to point along the positive z-direction - one then can boost the particle along the z-axis and rotate its momentum vector to achieve any other state of motion while preserving its helicity. The final state (a pair of spin-1/2 particles) may have any of a number of helicity quantum numbers available to them, depending on the angle of decay with respect to the z-axis and the spin vector of the parent. In the case depicted in Fig. 7, only one helicity state is allowed, and we can denote it $|\uparrow\downarrow\rangle$(or $|\downarrow\uparrow\rangle$if we swapped particle 1 and 2) in the product-ket notation. However, in general the decay angle could be non-zero, and all we know is the projection of $\vec{S}$ on the z-axis (that is the only definite thing we know about the orientation of $\vec{S}$), so any of a number of helicity states are possible. We can label them as product kets:

$$|m_1, \lambda_1; m_2, \lambda_2\rangle$$

and write them down (helicity is another two-state problem - so we already know how to represent them in ket space:

$$\begin{aligned} |1\rangle &= |+\uparrow;+\uparrow\rangle \\ |2\rangle\,,\ |3\rangle &= |+\uparrow;+\downarrow\rangle\ ,\ |+\downarrow;+\uparrow\rangle \\ |4\rangle &= |+\downarrow;+\downarrow\rangle\,. \end{aligned}$$

This lets us represent the final states with their helicity quantum numbers.

In general, if we want to describe what happens during the transition from an initial state to a set of final states, we have to do scattering theory and introduce the "S matrix," which tells us how the states evolve at all times. We would then compute an amplitude like so,

$$\mathcal{A} \propto \sum_n \langle f_n| S |i\rangle\,,$$

which is related to a physical observable by squaring the amplitude:

$$|\mathcal{A}|^2 = \mathcal{A}^*\mathcal{A}.$$

Since we are going to square the amplitude anyway, let's consider $\mathcal{A}^*$. In our specific case, we are interested in considering the decay of a particle at rest into a pair of particles at angle θ with respect to the z-axis, and about the y-axis at an angle ϕ. We can write our amplitude as

$$\mathcal{A}^*_{1,1} \propto \langle 1, 1| S |+, \lambda; +, \lambda'; \Omega(\theta, \phi)\rangle$$

Taking into account the spin projections on the z-axis, the helicity states, and the angular orientations of the final-state particles. However, we want to determine this angular dependence; so all we need to do is instead think about how we would rotate particles of these helicities away from $\theta = \phi = 0$ (basically, on the z-axis) to any other orientation. This is achieved by applying two successive rotations to

$$|+, \lambda; +, \lambda'; \Omega(0, 0)\rangle$$

that then rotate this state around the z-axis (by the Euler Angle ϕ) and then about the y-axis (by the Euler Angle θ). This takes the form:

$$U(\theta)U(\phi)\left|+,\lambda;+,\lambda';\Omega(0,0)\right\rangle = \left|+,\lambda;+,\lambda';\Omega(\phi,\theta)\right\rangle .$$

where we can write:

$$U(\theta)U(\phi) = U(\Omega).$$

Our amplitude becomes:

$$\begin{aligned}
\mathcal{A}^*_{1,1} &\propto \left\langle 1,1\right| S \left|+,\lambda;+,\lambda';\Omega(\phi,\theta)\right\rangle \\
&= \left\langle 1,1\right| S\, U(\Omega) \left|+,\lambda;+,\lambda';\Omega(0,0)\right\rangle \\
&= \sum_{j,m} \left\langle 1,1\right| S \left|j,m\right\rangle \left\langle j,m\right| U(\Omega) \left|+,\lambda;+,\lambda';\Omega(0,0)\right\rangle \\
&= \left\langle 1,1\right| S \left|1,1\right\rangle \left\langle 1,1\right| U(\Omega) \left|+,\lambda;+,\lambda';\Omega(0,0)\right\rangle .
\end{aligned}$$

The product-ket states have projections of their spin such that $m = \lambda_1 - \lambda_2$, which will then be rotated by the rotation matrix into the state $\langle 1,1|$ with its definite S_z value and definite helicity. This can be generically written as $\left\langle j',m'\right| U(\theta,\phi) \left|j,\lambda_1-\lambda_2\right\rangle$, and are given by the "Wigner d-functions," [26] which are also provided for you in Fig. 6. These have the form:

$$\left\langle j',m'\right| U(\Omega) \left|j,\lambda_1-\lambda_2\right\rangle = d(\theta)^j_{m',m=\lambda_1-\lambda_2} e^{-i(\lambda_1-\lambda_2)\phi}.$$

The exponential term (and thus dependence on the orientation about the z axis) will vanish when you compute each contribution to the total amplitude$\mathcal{A}^2 = \sum_n |\mathcal{A}_n|^2$; the other angular component, however, will remain.

We have three amplitudes that contribute to our sum:

$$\mathcal{A}^*_{1,1} \propto \left\langle 1,1\right| S \left|1,1,0\right\rangle \left\langle 1,1\right| d(\theta)^1_{1,1} e^{-i\phi}$$

so that

$$\mathcal{A}^*_{1,1} \propto \frac{1}{2}(1+\cos\theta)e^{-i\phi}.$$

Then we have two others:

$$\begin{aligned}
\left|1,0\right\rangle \longrightarrow \mathcal{A}^*_{1,0} &\propto -\frac{1}{\sqrt{2}}\sin\theta \\
\left|1,-1\right\rangle \longrightarrow \mathcal{A}^*_{1,-1} &\propto \frac{1}{2}(1-\cos\theta).
\end{aligned}$$

If one cannot distinguish in the final state (say, by applying an experimental appratus) the various helicity states of the fermions, then one has to sum over the final-state helicities ($m = 1, 0, -1$). The total amplitude will then be something like:

$$\mathcal{A}^2 \propto \epsilon_1(1+\cos\theta)^2 + \epsilon_2 \sin^2\theta + \epsilon_3(1-\cos\theta)^2. \tag{5}$$

We could choose some example coefficients and calculated the angular dependence (amplitude vs. $\cos\theta$, for instance). This is shown in Fig. 8. But will we see this in nature? The Stanford Large Detector (SLD) at the SLAC Laboratory produced Z bosons by colliding polarized electron and positron beams. The degree of polarization could be used to alter the coefficients of the spin admixture present in the spin-1 Z boson at the time of its production. We see that the structure predicted by the angular momentum conservation in the decay is realized in nature.

Of course, the details matter. The precise form of the coefficients and other kinematic effects are only predicted by considering the full scattering theory in relativistic quantum mechanics.

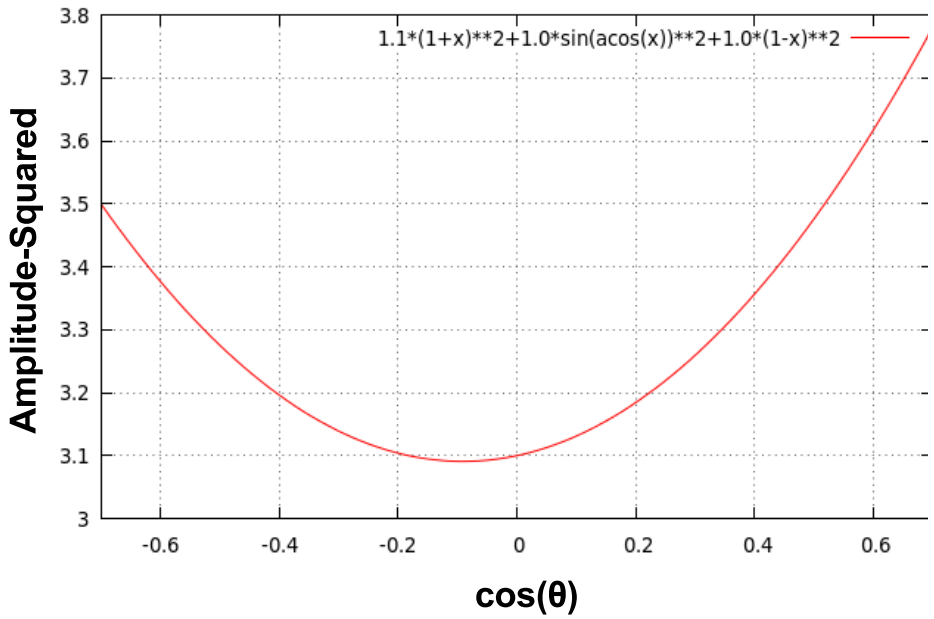


Figure 8: GNUPlot image of the amplitude-squared vs. $\cos\theta$ for an arbitrary set of coefficients in Equation 5.

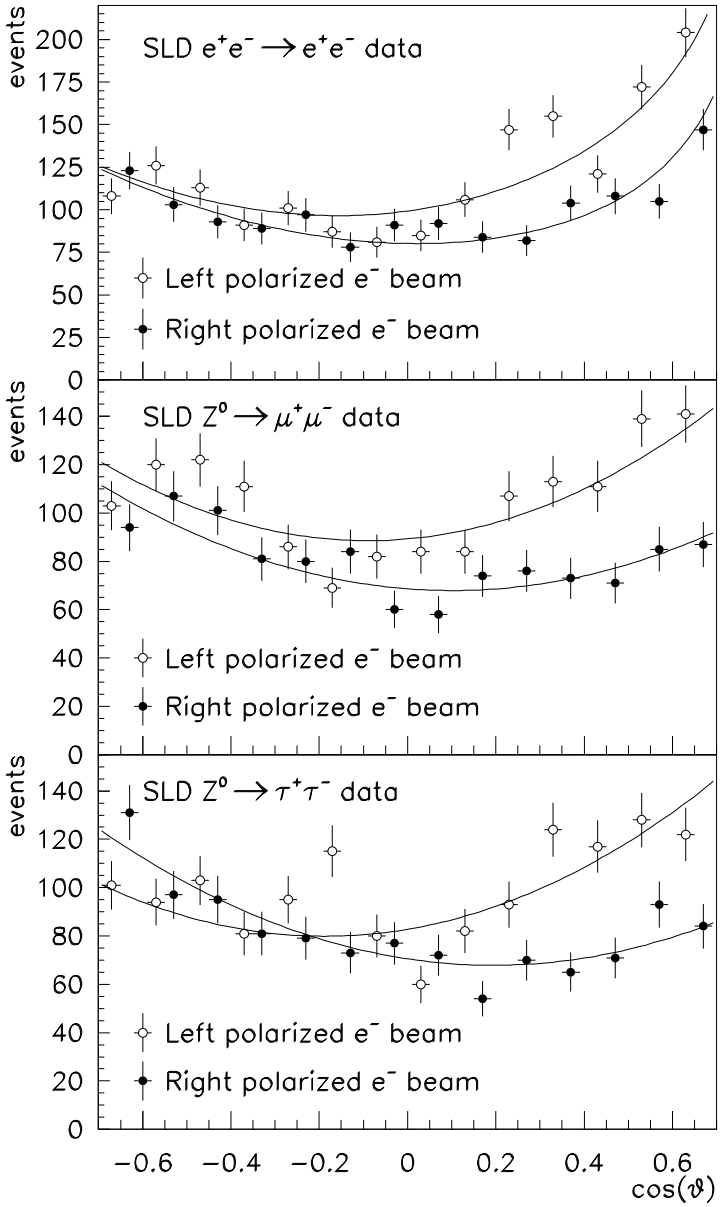


Figure 9: Polar angle distribution for Z^0 decays to e, μ and τ pairs for the 1994-5 SLD run. The asymmetries in the 1993 data look similar but are less pronounced due to the lower polarization. Figure and caption from Ref. [27].

10 Advanced Topic: Spin and Interactions

In this section of the course, we'll explore the implications for particles (with and without spin) interacting with one another. We'll build on what we've developed so far, but we need to add some more pieces to enhance our toolkit for describing interactions.

10.1 Non-Relativistic Perturbation Theory and Interactions

Consider the general problem of a free particle that comes into contact at some point with an external potential (e.g. experiences an "interaction" via the potential) and then continues onward. We can denote the initial state by labeling it i and the final state by labeling it f .

Consider the purely free-particle situation for a second, we can write the Schroedinger Equation as:

$$H_0\phi_n = E_n\phi_n$$

where H_0 is the free-particle Hamiltonian ($\frac{-\hbar^2}{2m}\nabla^2$) and the ϕ_n are the eigenstates of the system (a single particle). We consider all of this happening in some spatial volume, V, and that the eigenstates are normalized and orthogonal:

$$\begin{aligned}\int_V \phi_n^*\phi_n \, d^3x &= 1 \\ \int_V \phi_m^*\phi_n \, d^3x &= 0 \; (m \neq n).\end{aligned}$$

These can be summarized in a single equation:

$$\int_V \phi_m^*\phi_n \, d^3x = \delta_{mn}.$$

The goal, of course, is to solve Schroedinger's Equation now with the particle in the presence of a potential, $V(\vec{x}, t)$:

$$(H_0 + V(\vec{x}, t))\psi = i\frac{d\psi}{dt}.$$

If we can find the eigenstates of the system, including the potential, we should be able to write each in the space-and-time separable form:

$$\phi_n(\vec{x}, t) = \phi(\vec{x})e^{-iE_nt}.$$

Any general solution of the SWE can be written in terms of these orthogonal solutions:

$$\psi = \sum_n a_n(t)\phi_n(\vec{x})e^{-iE_nt}.$$

We don't know the solutions, but we can try to sort out these coefficients, $a_n(t)$. We insert the above solution into the SWE:

$$i\frac{d}{dt}\left[\sum_n a_n(t)\phi_n(\vec{x})e^{-iE_nt}\right] = \sum_n H_0\left[a_n(t)\phi_n(\vec{x})e^{-iE_nt}\right] + \sum_n V(\vec{x}, t)\left[a_n(t)\phi_n(\vec{x})e^{-iE_nt}\right].$$

We know that $H_0\phi_n(\vec{x}) = E_n\phi_n(\vec{x})$, so that:

$$\sum_n H_0\left[a_n(t)\phi_n(\vec{x})e^{-iE_nt}\right] = \sum_n (E_n)a_n(t)\phi_n(\vec{x})e^{-iE_nt}$$

The free-particle Hamiltonian affects no other parts of the wave function, term-by-term. We can also write the left side by applying the chain rule:

$$\begin{aligned} i\frac{d}{dt}\left[\sum_n a_n(t)\phi_n(\vec{x})e^{-iE_nt}\right] &= i\left[\sum_n \phi_n(\vec{x})\left(e^{-iE_nt}\frac{d}{dt}a_n(t) + a_n(t)\left((-iE_n)\,e^{-iE_nt}\right)\right)\right] \\ &= i\left[\sum_n \phi_n(\vec{x})e^{-iE_nt}\left(\frac{d}{dt}a_n(t) - ia_n(t)E_n\right)\right] \\ &= i\quad\sum_n \phi_n(\vec{x})e^{-iE_nt}\frac{d}{dt}a_n(t) + \sum_n E_n a_n(t)\phi_n(\vec{x})e^{-iE_nt} \end{aligned}$$

We see that some terms cancel on the left and right of the SWE, leaving:

$$i\sum_n \phi_n(\vec{x})e^{-iE_nt}\frac{d}{dt}a_n(t) = \sum_n V(\vec{x},t)\left[a_n(t)\phi_n(\vec{x})e^{-iE_nt}\right].$$

We want an equation for just the coefficients, and the way we can achieve this is to take into account the orthonormality of the eigenstates. Multiplying from the left by $\phi_f^*(\vec{x})$:

$$i\sum_n \phi_f^*(\vec{x})\phi_n(\vec{x})e^{-iE_nt}\frac{d}{dt}a_n(t) = \sum_n \phi_f^*(\vec{x})V(\vec{x},t)\left[a_n(t)\phi_n(\vec{x})e^{-iE_nt}\right]$$

and then integrating over the volume, leads to:

$$i\int_V \phi_f^*(\vec{x})\phi_n(\vec{x})e^{-iE_nt}\frac{d}{dt}a_n(t)\,d^3x = \int_v \sum_n \phi_f^*(\vec{x})V(\vec{x},t)\left[a_n(t)\phi_n(\vec{x})e^{-iE_nt}\right]\,d^3x$$

which yields:

$$i\delta_{fn}e^{-iE_nt}\frac{d}{dt}a_n(t) = ie^{-iE_ft}\frac{d}{dt}a_f(t) = \int_v \sum_n \phi_f^*(\vec{x})V(\vec{x},t)\left[a_n(t)\phi_n(\vec{x})e^{-iE_nt}\right]\,d^3x.$$

Finally, we move everything except the time-derivative to the right-hand side of the equation and get what we want:

$$\frac{da_f(t)}{dt} = -i\sum_n a_n(t)\int_V \phi_f^*(\vec{x})V(\vec{x},t)\phi_n(\vec{x})e^{-i(E_n-E_f)t}\,d^3x.$$

We now have an equation that relates the time-dependence of a single coefficient, $a_f(t)$, to a sum over all other coefficients times integrals of the probability density of an eigenstate interacting with the potential and yielding a projection along the eigenstate $\phi_f(\vec{x})$.

Let us then try to go a little further. Let us simplify the problem by imagining that, at first, we prepare the free particle system in an eigenstate of the SWE. Let us choose $n = i$, where i denotes the initial eigenstate (one of many from which we could have chosen) in which we prepare the system. This then forces $a_i = 1$ and $a_m = 0$ for $m \neq i$. Our equation then becomes:

$$\frac{da_f}{dt} = -i\int_V \phi_f^*(\vec{x})V(\vec{x},t)\phi_i(\vec{x})e^{-i(E_i-E_f)t}\,d^3x.$$

Let us further imagine that over a period of time, T, the free-particle travels, then interacts via $V(\vec{x},t)$, and then exits in its final-state. We can consider the case of a "nearly free particles" as a small "perturbation" to the purely free-particle scenario. That is, consider a time spent in the potential that is:

- transient - the time period spent inside the potential, $\delta t << T$.

- weak - the potential strength itself is very small in magnitude

We can then see what happens. Imagine that the potential interaction occurs as time $t = 0$, with $\delta t << T$. The particle is sent in for the interaction at time $t = -T/2$ and concludes its journey in its final state at time $t = T/2$. This allows us to assume that the initial conditions essentially hold at all times, since the influence of the potential will be small. We can find the coefficient at any time t by integrating the above equation from $-T/2$ to an arbitary time, t , where $t \leq T/2$:

$$a_n = -i \int_{-T/2}^{t} \int_V \phi_f^*(\vec{x}) V(\vec{x}, t) \phi_i(\vec{x}) e^{-i(E_i - E_f)t} \, d^3x \, dt.$$

We can identify the space-time volume element, $d^4x = d^3x \, dt$. This is convenient for later. We are specifically interested in the coefficient at the time $T/2$, when the interaction has long-ceased. This specific value of the coefficient is denoted:

$$T_{fi} \equiv a_f(T/2) = -i \int_{-T/2}^{+T/2} \int_V \left[\phi_f^*(\vec{x}) e^{iE_f t}\right] V(\vec{x}, t) \left[\phi_i(\vec{x}) e^{-iE_i t}\right] \, d^3x \, dt.$$

This represents an integral over all relevant space and time points in the problem. We can then write the above coefficient in the compact format:

$$T_{fi} = -i \int \left[\psi_f^*(\vec{x}, t) V(\vec{x}, t) \psi_i(\vec{x}, t)\right] \, d^4x.$$

This is only a good approximation if $a_f << 1$, which was assumed in the above calculations (we can check the correction to this assumption later).

You can show that the above quantity, while temptingly assumed to be related to the probability for the transition from the initial state to the final state, is not physically meaningful as such. However, a slight redefinition of the above yields the physically meaningful quantity.

If we consider a time-independent potential, $V(\vec{x}, t) \to V(\vec{x})$, we can write:

$$T_{fi} = -i \left[\int \phi_f^* V \phi_i \, d^3x\right] \left[\int_{-\infty}^{+\infty} e^{-i(E_i - E_f)t} \, dt\right].$$

The time integral is just the Dirac Delta Function:

$$\delta(a - b) \equiv \frac{1}{2\pi} \int e^{ix(a-b)} dx$$

so that

$$T_{fi} = -2\pi i \left[\int \phi_f^* V \phi_i \, d^3x\right] \delta(E_f - E_i).$$

we can then define:

$$W = \lim_{T \to \infty} \frac{|T_{fi}|^2}{T}$$

which is the probability per unit time of transitioning from the initial to the final state via this potential. Recalling that $|T_{fi}|^2 = T_{fi}^* T_{fi}$, defining $V_{fi} = \int \phi_f^* V \phi_i \, d^4x$, and remembering that the square of a delta function is just a delta function,

$$W = \lim_{T \to \infty} \left[\frac{1}{T}\right]$$

11 The Helicity Formalism (Roberto Vega)

This part of the lecture is contributed by Prof. Roberto Vega, and based on his notes and lecture.

Consider a particle, labelled α, decaying into two particles, 1 and 2. If you know the spin of the parent, and you know the spin of the final-state particles - say, $S_1 = 0$ and $S_2 = \frac{1}{2}$ - then already that information lets you know about the outgoing angular distribution of particles will look like. This is true even without knowing exactly the Hamiltonian. You need only know the angular momentum, that it's conserved, and you can proceed.

Generally speaking, for this case, you get a superposition of terms like

$$dY_1^0\chi_{1/2,1/2} + cY_1^{-1}\chi_{1/2,-1/2} + bY_0^0\chi_{1/2,1/2}.$$

You can then determine from these functions the possible angular distributions. This is the common procedure.

The advantage of this procedure is that the angular momentum is defined in the rest frame of the parent. However, you also know the spin states in the rest frame of the daughters. You have to shuffle between the systems by rotation. This is tedious, however. If you want to consider polarization effects, this is not a convenient method.

This is why we use "the helicity formalism." It is based on the observation that the helicity operator, the projection of spin in the direction of motion,

$$\vec{S}\cdot\hat{p} = \vec{J}\cdot\hat{p},$$

commutes with the momentum and the total angular momentum. That is, since the above relation holds,

$$\begin{array}{ccc} [\vec{P}, \vec{J}\cdot\hat{p}] = 0 & \text{and} & [J_i, \vec{J}\cdot\hat{p}] = 0 \\ & \text{while} & \\ [P_i, J_j] \neq 0 & \text{and} & [\vec{P}^2, J_i] \neq 0. \end{array}$$

The eigenvalue associated with the operator is what is meant by "helicity." Since we can choose any direction to be the z-axis along which we quantize angular momentum projection, it is convenient to choose the direction of motion, $\hat{p} = \hat{z}$.

Thanks to commutation, we see that we have two independent sets of basis vectors. One set can be labeled with the eigenvalues of the momentum operator, $\vec{P}$, and by the eigenvalues of the helicity operator. These are "plane-wave helicity states":

$$|p, \theta, \phi; \lambda\rangle\,.$$

The other, we can label with the eigenvalues $|\vec{P}|$, J^2, J_z, and the helicity operator:

$$|j, m, |p|; \lambda\rangle\,.$$

This is the "spherical wave helicity basis." Also, if we consider the final-state particles from the decay example above, each is also describable in its own helicity basis.

The helicity formalism is one that labels the states using either of these two basis sets. Why is this advantageous? Because the helicity operator is invariant under Lorentz boosts along the direction of motion. We can work in any frame; for instance, we can choose to start in the center-of-mass frame, and then boost, and out results still describe the physics in other frames. This formalism applies equally well to both massive and massless particles. Helicity is invariant for massive and massless particles. In the other approach - labeling by the spin and orbital angular momentum - we have to treat massless particles specially (they cannot be at rest - we cannot use the rest frame). In the Helicity Formalism, you can deal with total angular momentum, and not worry about the details of orbital angular momentum. As a bonus, helicity amplitudes for scattering give us a way of easily writing down the polarization effects in a scattering system.

What is the general idea? The idea is that we have a scattering:

$$\alpha \to 1 + 2.$$

Here, α has a well-defined J and M - the total angular momentum and its projection in the z-direction. This is just a decay. We can write the amplitude for this

$$\mathcal{A} = \langle \vec{p}_1 \vec{p}_2; \lambda_1 \lambda_2 | T | JM \rangle,$$

where $T = e^{-(i/\hbar)Ht}$ is an operator that evolves the states overall times (see the earlier section on scattering). T is a scalar,

$$[T, J_i] = [H, J_i] = 0,$$

since T is just a series whose terms contain powers of the Hamiltonian, and the Hamiltonian commutes with the angular momentum operator for rotationally invariant systems (which is what we are considering througout these notes - systems where angular momentum is conserved and which are thus rotationally invariant). We can then expand the above amplitude in any other basis we like. For instance, we can write it in the basis of total-j states:

$$\mathcal{A} = \sum_{j',m'} \langle \vec{p}_1 \vec{p}_2; \lambda_1 \lambda_2 | j' m' \rangle \langle j' m' | T | JM \rangle.$$

Since T is a scalar, we can write this matrix element using the Wigner-Eckhart Theorem:

$$\langle j' m' | T | JM \rangle = \langle j' m' | 00 JM \rangle \langle |\vec{p}|, j, \lambda | T | J \rangle = \delta_{Mm',Jj'} \langle |\vec{p}|, j, \lambda | T | J \rangle \equiv T_{\lambda_1 \lambda_2}.$$

which has just a Clebsch-Gordon Coefficient relating the two bases to one another. This Clebsch-Gordon Coefficient in this specific case is just a delta function - it vanishes if $m' \neq M$ and $j' \neq J$. It is also independent of the index, m. We can label it any way we like. The other term:

$$\langle \vec{p}_1 \vec{p}_2; \lambda_1 \lambda_2 | j' m' \rangle \equiv D^J_{M\lambda}(\theta, \phi).$$

So in the end, we just have:

$$\mathcal{A} = D^J_{M\lambda}(\theta, \phi) T_{\lambda_1 \lambda_2}.$$

where $\lambda = \lambda_1 - \lambda_2$.

The hard work of determining the angular dependence of the final state is let to determining these "D-matrices," which were partially introduced earlier and now will be defined more carefully.

11.1 Massless and Massive Particles

For a massive particle,

$$\lambda = s, s-1, ..., -s.$$

The helicity states for a complete basis:

$$|\vec{p}, \lambda\rangle .$$

For a massless particle, you can still define the helicity basis. But there is a difference. Here,

$$\lambda = \pm s.$$

The nature of spin for a massless particle is quite different for a massless particle. Massive particles have spin due to the fact that rotations occur in the group SU(2). But for massless particles, rotations involve three parameters - the three Euler Angles, if you like.

Remember that boosts along the direction of motion won't change helicity. Boosts in any other direction, of course, will change the helicity.

11.2 Rotations - Revisited

The rotation operator that acts on kets is unitary, and in general depends on three parameters:

$$U(R_{\hat{n}}(\varphi)) = e^{-i\varphi\hat{n}\cdot\vec{J}}$$

in natural units. This is a rotation by φ about $\vec{J}$. The three parameters are (1) the rotation angle, and (2,3) the two parameters that define $\hat{n}$ - the polar and azimuthal angles. We can rotate a state like so:

$$U(R)\,|\psi\rangle = |\psi'\rangle\,.$$

These matrices act on other operators via:

$$UAU^{\dagger} = A'.$$

For example, for spin-1/2:

$$\vec{J} = \frac{1}{2}\vec{\sigma},$$

we can rewrite the exponential function in terms of sines and cosines and find:

$$U(R_{\hat{n}}(\varphi)) = \cos(\varphi/2) - i\vec{\sigma}\cdot\hat{n}\sin(\varphi/2).$$

The properties of the dot product in the above are:

$$\begin{aligned}(\hat{n}\cdot\vec{\sigma})^2 &= \hat{n}^2 = 1\\ (\hat{n}\cdot\vec{\sigma})^3 &= \hat{n}\cdot\vec{\sigma}.\end{aligned}$$

For spin-1:

$$U = 1 + (\hat{n}\cdot\vec{J})^2(\cos\varphi - 1) - i(\hat{n}\cdot\vec{J})\sin\varphi.$$

We know that:

$$\sigma_z\,|1/2, 1/2\rangle = +\,|1/2, 1/2\rangle$$

which can be written:

$$\hat{z}\cdot\vec{\sigma}\,|1/2,\sigma\rangle = \sigma\,|1/2,\sigma\rangle\,.$$

We can then rotate:

$$U\hat{z}\cdot\vec{\sigma}U^{\dagger}U\,|1/2,\sigma\rangle = \sigma U\,|1/2,\sigma\rangle$$

and one can then write this out in terms of rotation angles and show that it does, in fact, behave exactly as you'd expect a rotation operator to behave.

Let's look at this using Euler Angles - first rotate about a selected axis, then away from the axis (around another one), and finally again about the original axis. In quantum mechanics, we can write this in a simple way:

$$U(\alpha,\beta,\gamma) = e^{-i\alpha J_z}e^{-i\beta J_y}e^{-i\gamma J_z}.$$

It's amazing that we can define an arbitrary rotation to a new point using these three simple rotations about just two axes - but it does work.

If S is a rotation, and I take another rotation, R, and calculate this:

$$SR(\vec{\psi})S^{-1} = R(S\vec{\psi}).$$

In other words, the rotation on the left is the same as just rotating the vector, $\vec{\psi}$, first by the matrix S and then defining R about the new axis. This can be proven, but we will merely employ this theorem.

We can then write:

$$\begin{aligned} U(\vec{\psi}\cdot\vec{J})^N U^\dagger &= (U(\vec{\psi}\cdot\vec{J})U^\dagger)^N \\ &= \psi_i U_\zeta J_i U_\zeta^\dagger = \psi_i S_{ij}^{-1} J_j = \psi_i' J_i. \end{aligned}$$

11.3 The Wigner D-Matrices

The Wigner D-matrices can now merely be written in terms of these rotations. We can define these matrices as mentioned earlier:

$$D^{(j)}_{m'm}(\alpha,\beta,\gamma) = \langle j'm'|\, U(\alpha,\beta,\gamma)\,|jm\rangle\,.$$

This works because $J^2$is a scalar (we are dealing with rotationally invariant systems!); that is:

$$UJ^2U^\dagger = J^2.$$

The above is not true for only J_z. While total angular momentum is conserved, components can be changed under a rotation. We can then proceed:

$$D^{(j)}_{mm'}(\alpha,\beta,\gamma) = e^{-i\alpha m'}e^{-i\gamma m}\,\langle j'm'|\, e^{-i\beta J_y}\,|jm\rangle\,.$$

where:

$$d^j_{m'm} \equiv \langle j'm'|\, e^{-i\beta J_y}\,|jm\rangle\,.$$

Note that:

$$\langle j'm'|\, e^{-i\pi J_y}\,|jm\rangle = (-1)^{j-m}\delta_{m,-m'}.$$

We will use this (though you can prove it from the rotation operator, if you choose).

For the plane-wave helicity states, we can write:

$$|p,\lambda\rangle\,.$$

We will define the *standard ket* as:

$$|p_z\lambda\rangle$$

with all momentum along the z-axis. We can then do everything relative to this *standard ket*. For a massive particle, the standard ket is $|0\lambda\rangle$ –at rest, with the spin pointing in the z-direction. For a massless particle, our convention will be to write the standard ket as $|\kappa_z\lambda\rangle$, where $\kappa^2 = 0$. From these standard kets, I can generate ANY state of momentum by applying a Lorentz Transformation for the massive case; for the massless case, I have to consider what happens to the particle under a parity transformation (since I can never "boost ahead" of the particle and thus reverse its direction of motion).

For the standard ket, we can start with writing all components of momentum explicitly:

$$|p,0,0;\lambda\rangle\,.$$

I can then rotate this momentum state in any direction to get the general problem of a particle with momentum $\vec{p} = (p_x,p_y,p_z)$, where the momentum may not lie entirely along one selected axis.

Let's rotate. Not that rotating about the z-direction does nothing to the state. We have to choose a "phase convention" for how to handle that rotation, since without a convention its effect is undefined. We choose the "Jacob-Wick Phase Convention" where we choose the phase angle $\varphi = \phi$. Rotating the standard ket:

$$U(\phi, \theta, \phi) \left| p, 0, 0; \lambda \right\rangle = \left| p, \theta, \varphi; \lambda \right\rangle ,$$

we can get a hint of how this is, in the end, merely doing to be related to the D-functions. We can write our standard ket in terms of the basis states of total-j:

$$\left| p, 0, 0; \lambda \right\rangle = \sum_{j,m} C_j^m \left| p, j, m; \lambda \right\rangle .$$

The λ that appears here is th esame on both sides. The only term that survives in the sum is then the one with $m = \lambda$:

$$\begin{aligned}
\left| p, 0, 0; \lambda \right\rangle &= \sum_j C_j \left| p, j, \lambda; \lambda \right\rangle \\
\left| p, \theta, \phi; \lambda \right\rangle &= \sum_j C_j U(\phi, \theta, \phi) \left| p, j, \lambda; \lambda \right\rangle \\
&= \sum_j C_j \sum_{j'm'} \left| p, j; , m'; \lambda \right\rangle \left\langle p, j', m'; \lambda \right| U(\phi, \theta, \phi) \left| p, j, m; \lambda \right\rangle \\
&= \sum_j C_j \sum_{j'm'} \left| p, j; , m'; \lambda \right\rangle D^{(j)}_{m'\lambda}(\phi, \theta, \phi).
\end{aligned}$$

We cannot change the magnitude of $\vec{J}$ in these transformations, so the only terms that survive are those where $j' = j$. Also, it must be that $m = m'$ - otherwise, the inner product vanishes. Thus:

$$\left| p, \theta, \phi; \lambda \right\rangle = \sum_{j,m} C_j \left| p, j, m; \lambda \right\rangle D^{(j)}_{m\lambda}(\phi, \theta, \phi).$$

If I then take the inner product:

$$\langle p, j', m'; \lambda | p, \theta, \phi, \lambda \rangle = \sum_{j,m} C_j D^{(j)}_{m\lambda}(\phi, \theta, \phi) \langle p, j', m'; \lambda | p, j, m; \lambda \rangle = C_j D^{(j)}_{m\lambda}(\phi, \theta, \phi).$$

It can be shown that these D-functions, for integer values of $j = \ell$, are proportional to the spherical harmonics, Y^ℓ_m.

If we then require that:

$$\langle p, \theta, \phi, \lambda | p, \theta, \phi, \lambda \rangle = 1,$$

it can be shown that:

$$C_j = \sqrt{\frac{2j+1}{4\pi}}.$$

This allows us to write down the angular distributions for a decay like $\alpha \to 1 + 2$. In that case, we just have two particles, back-to-back in the center-of-mass frame of the parent. The net helicity of the system is $\lambda = \lambda_1 - \lambda_2$. One only still needs two angles, θ and ϕ.

11.4 Helicity States under Transformations

We now want to explore what happens to helicity states under parity and reflection transforms. This is useful; once we know these relations, we can simply emply them to write states that relate to such transformations.

References

[1] P. Zeeman, "On the influence of magnetism on the nature of the light emitted by a substance," *Phil. Mag.*, vol. 43, p. 226, 1897.

[2] P. Zeeman, "Doubles and triplets in the spectrum produced by external magnetic forces," *Phil. Mag.*, vol. 44, p. 266, 1897.

[3] P. Zeeman, "The effect of magnetisation on the nature of light emitted by a substance," *Nature*, vol. 55, p. 1424, 1897.

[4] *The Astrophysical Journal*, vol. 49, pp. 153–178.

[5] A. Einstein, "Uber die von der molekularkinetischen theorie der warme geforderte bewegung von in ruhenden flÃŒssigkeiten suspendierten teilchen," *Annalen der Physik*, vol. 322, pp. 549–560, 1905.

[6] J. Thompson, "Cathode rays," *The Electrician*, vol. 39, p. 104.

[7] J. Thompson, "Cathode rays," *Philosophical Magazine*, vol. 44, p. 293.

[8] W. Foundation, "Plum pudding model," *Wikipedia*, vol. https://en.wikipedia.org/wiki/Plum_pudding_model, 2014.

[9] W. Foundation, "Rutherford model," *Wikipedia*, vol. https://en.wikipedia.org/wiki/Rutherford_model, 2014.

[10] W. Foundation, "Bohr model," *Wikipedia*, vol. https://en.wikipedia.org/wiki/Bohr_model, 2014.

[11] E. Rutherford, *The Scattering of α and β Particles by Matter and the Structure of the Atom*, vol. 21 of *6*. 1911.

[12] N. Bohr, *On the Constitution of Atoms and Molecules, Part I*, vol. 26 of *151*. 1913.

[13] N. Bohr, *On the Constitution of Atoms and Molecules, Part II: Systems Containing Only a Single Nucleus*, vol. 26 of *153*. 1913.

[14] N. Bohr, *On the Constitution of Atoms and Molecules, Part III: Systems containing several nuclei*, vol. 26. 1913.

[15] N. Bohr, "The spectra of helium and hydrogen," *Nature*, vol. 92 (2295), pp. 857–875, 1914.

[16] N. Bohr, "Atomic structure," *Nature*, vol. 107 (2682), pp. 104–107, 1921.

[17] A. Sommerfeld, *Atombau und Spektrallinien'*. Braunschweig: Friedrich Vieweg und Sohn, 1919.

[18] W. Gerlach and O. Stern, "Das magnetische moment des silberatoms," *Zeitschrift fur Physik*, vol. 9, pp. 353–355, 1922.

[19] F. Bretislav and H. Dudley, "Stern and gerlach: How a bad cigar helped reorient atomic physics," *Physics Today*, vol. 56(12), p. 53, 2003.

[20] P. Dirac, "Relativistic quantum mechanics," *Proceedings of the Royal Society A*, vol. 136, 1932.

[21] W. Foundation, "Cartesian coordinate system," *Wikipedia*, vol. https://en.wikipedia.org/wiki/Cartesian_coordinate_sys 2014.

[22] W. Foundation, "Spherical coordinate system," *Wikipedia*, vol. https://en.wikipedia.org/wiki/Spherical_coordinate_syst 2014.

[23] W. Foundation, "Spherical harmonics," *Wikipedia*, vol. https://en.wikipedia.org/wiki/Spherical_harmonic, 2014.

[24] B. Barish and R. Stroynowski, "The Physics of the tau Lepton," *Phys.Rept.*, vol. 157, p. 1, 1988.

[25] J. Beringer *et al.*, "Review of Particle Physics (RPP)," *Phys.Rev.*, vol. D86, p. 010001, 2012.

[26] D. Pendleton, "Euler angle geometry, helicity basis vectors, and the Wigner D-function addition theorem," *American Journal of Physics*, vol. 71, pp. 1280–1291, 2003.

[27] K. Abe *et al.*, "Direct measurement of leptonic coupling asymmetries with polarized Zs," *Phys.Rev.Lett.*, vol. 79, pp. 804–808, 1997.